ESAIM: COCV 28 (2022) 77
https://doi.org/10.1051/cocv/2022066

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

SECOND-ORDER ANALYSIS OF FOKKER–PLANCK ENSEMBLE OPTIMAL CONTROL PROBLEMS*

Jacob Körner** and Alfio Borzì

Abstract. Ensemble optimal control problems governed by a Fokker–Planck equation with space-time dependent controls are investigated. These problems require the minimisation of objective functionals of probability type and aim at determining robust control mechanisms for the ensemble of trajectories of the stochastic system defining the Fokker–Planck model. In this work, existence of optimal controls is proved and a detailed analysis of their characterization by first- and second-order optimality conditions is presented. For this purpose, the well-posedness of the Fokker–Planck equation, and new estimates concerning an inhomogeneous Fokker–Planck model are discussed, which are essential to prove the necessary regularity and compactness of the control-to-state ma p appearing in the first-and second-order analysis.

Mathematics Subject Classification. 35Q84, 35Q93, 49J20, 49K20.

Received October 22, 2021. Accepted October 8, 2022.

1. Introduction

This paper is devoted to the analysis of problems of optimal control of ensembles of trajectories of stochastic drift-diffusion models from a statistical perspective. The notion of ensemble controls was proposed by R.W. Brockett in [14–16], while considering different feedback control strategies and the corresponding trade-off in implementation and performance. In this statistical approach, the Fokker–Planck (FP) equation, corresponding to the given stochastic model, governs the evolution of the entire ensemble of trajectories of the model (with a distribution of initial conditions), and an expected-value cost functional accommodates all possible stochastic realizations, thus allowing the design of closed-loop control mechanisms.

Recently, Brockett's research programme has received much impetus through novel theoretical and numerical work focusing on deterministic models with random initial conditions and the corresponding Liouville equation [5, 6], and in the case of a linear Boltzmann equation [7]. The modelling and simulation of FP ensemble optimal control problems has been investigated in view of their large applicability [12, 34–36]. However, in comparison to the amount of work on FP control problems with quadratic objectives [2, 11, 23], much less effort has been put in the analysis of FP ensemble optimal control problems. On the other hand, in the former case, only very recently a detailed investigation of second-order optimality conditions of FP problems with time-dependent

*Partially supported by the BMBF-Project iDeLIVER: Intelligent MR Diagnosis of the Liver by Linking Model and Data-driven Processes.

Keywords and phrases: Optimal control theory, Fokker-Planck equation, ensemble optimal control problems, second-order analysis, optimality conditions, stochastic drift-diffusion process

Institut für Mathematik, Universität Würzburg, Emil-Fischer-Strasse 30, Würzburg 97074, Germany.

** Corresponding author: jacob.koerner@mathematik.uni-wuerzburg.de

controls has been presented by M.S. Aronna and F. Tröltzsch [4]. The motivation for this analysis is manifold, as illustrated in, *e.g.*, [19, 20], and essential for exploring, in particular, stability and approximation issues. For this reason, we would like to contribute to the further advance of FP ensemble optimal control problems with the second-order analysis presented in this paper.

Fokker–Planck ensemble optimal control problems represent a new framework for designing robust controls of stochastic models that appear in different scientific disciplines. In the field of epidemic problems, stochastic models are of great interest because they take into account the fact that biological actions and human behaviour are subject to random fluctuations [25, 26, 42]. Stochastic systems also appear in the modelling and control of collective motion in biological systems – such as the movement of groups of bacteria or herds of animals – as well as in pedestrian motion and traffic flows [10, 34–36]. In these and many other works, the focus is on drift-diffusion models as follows

$$\mathrm{d}X_t = B[u](t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t, \quad X_{|t=0} = X_0, \qquad t \geqslant 0, \tag{1.1}$$

where $X_t \in \mathbb{R}^n$ is a continuous-time stochastic process, $\mathrm{d}W_t \in \mathbb{R}^n$ a Wiener process, and X_0 a given n-dimensional random variable. With $B[u]$ we denote the drift including a control mechanism, and σ represents a variable dispersion matrix.

In the framework of stochastic control theory [9], the objective to be minimized subject to (1.1) is the following expected-value functional

$$J(X, u) = \mathbb{E}\left[\int_0^T \mathcal{R}\big((t, X_t, u(t, X_t)\big)\,\mathrm{d}t + \mathcal{T}(X_T)\right], \tag{1.2}$$

where $\mathcal{R}[u] = \mathcal{R}[u](t, x)$ is usually called the running cost, and $\mathcal{T} = \mathcal{T}(x)$ is referred to as the terminal observation.

We remark that the state of a stochastic process can be completely characterized by the shape of its statistical distribution, which is represented by the corresponding probability density function (PDF) that we denote with $p = p(t, x)$. Further, we have a fundamental result in statistical mechanics showing that the evolution of the PDF associated to X_t is modelled by the following FP problem

$$\partial_t p + \operatorname{div}\big(B[u]\,p\big) - \sum_{i,j=1}^{n} \partial^2_{x_i x_j}\big(a_{ij}\,p\big) = 0, \tag{1.3}$$

$$p(0, x) = p_0(x),$$

$$F(t, x)\cdot \hat{n}(x) = 0. \tag{1.4}$$

In this problem, the FP equation (1.3) has an advection term that corresponds to the drift and the diffusion matrix $a = \frac{1}{2}\sigma^{\intercal}\sigma$. The initial data is given by the initial PDF p_0 of X_0. Further, we assume that the process is bounded and conserved in a bounded domain $\Omega \subset \mathbb{R}^n$, which results in the flux-zero boundary conditions (1.4), where the vector probability density flux F has the components

$$F_j(t, x) := \sum_{i=1}^{n} \partial_{x_i}\big(a_{ij}(t, x)\,p(t, x)\big) - \big(B[u]_j(t, x)\,p(t, x)\big), \qquad j = 1, \dots, n,$$

and $\hat{n}$ denotes the outward normal unit vector on $\partial\Omega$. We can see that in passing from trajectories to PDFs, the space of the state X has become the space of an independent variable x with the same dimension.

Our analysis focuses on the following cost functional

$$J(p,u) = \int_\Omega \int_0^T \mathcal{R}[u](t,x)\, p(t,x)\, \mathrm{d}t\, \mathrm{d}x + \int_\Omega \mathcal{T}(x)\, p(T,x)\, \mathrm{d}x + \frac{\gamma}{2}\|u\|^2_{H^1(\Omega_T)}. \tag{1.5}$$

Notice that the first two terms in $J(p,u)$ correspond to (1.2) and define, for an appropriate choice of $R[u]$, an ensemble cost functional as discussed by Brockett. However, we have added an additional term corresponding to a H^1-cost of the control that is not subject to averaging. This additional term plays a crucial role in our investigation, while in the last section of this paper, we consider variants of (1.5) with $\gamma = 0$ that accommodate different Brockett's ensemble functionals.

Our FP ensemble optimal control problems are formulated as the minimisation of (1.5) subject to the differential constraint given by (1.3)–(1.4), and space-time dependent controls with admissible values in convex compact sets. While we do not investigate the case of additional constraints on the state p or the control u, for these cases we refer to [38, 39] for a first-order analysis devoted to parabolic and hyperbolic models, and to [30] for the case of a semilinear elliptic equation.

Our analysis of the Fokker–Planck ensemble optimal control problem follows the work in [4], where a detailed first- and second-order analysis for time dependent n-dimensional L^2-controls $u = u(t)$ and quadratic objectives is presented. However, in our ensemble setting, we introduce a different functional framework that allows to accommodate space-time dependent controls and makes possible to perform second-order analysis. The main difficulty of the analysis of the optimal control problem (1.3)–(1.5) is treating the flux term $\operatorname{div}\left(B[u](t,x)\, p(t,x)\right)$ due to its bilinear form and the control being subject to differentiation, which leads to results with lower regularity for the control-to-state map.

The paper is organized as follows. The next section discusses existence and regularity of weak solutions to our Fokker–Planck problem. We also point out the properties of these solutions as PDF functions. In Section 3, in view of our first- and second-order analysis, we discuss a related inhomogeneous FP equation and present L^∞-estimates concerning its solution.

In Section 4, we define and investigate the FP control-to-state operator and show its continuity, Fréchet differentiability and compactness on proper sets. In Section 5, we introduce the reduced cost functional and prove existence of optimal controls in the given admissible control sets. In Section 6, we discuss the characterization of these optimal controls by first-order optimality conditions. For this purpose, we introduce and analyze a FP adjoint equation and discuss the Fréchet derivative of the reduced cost functional in order to formulate the first-order necessary optimality conditions. Section 7 is devoted to the analysis of necessary and sufficient second-order optimality conditions for our FP ensemble optimal control problems. In Section 8, we discuss extension of our results to FP control problems with variants of Brockett's ensemble cost functional. These extension accommodate different settings that have appeared in the scientific literature, including Brockett's approach to minimal attention feedback control. A section of conclusion completes this work.

Notation and general assumptions

Let $n \in \mathbb{N} = \mathbb{N}\backslash\{0\}$ denote the dimension of x and let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a Lipschitz boundary $\partial\Omega$. For arbitrary but fixed $T > 0$, we define $\Omega_T :=]0,T[\, \times \Omega$ and denote for a normed space Y and $1 \leqslant p \leqslant \infty$ the Bochner-spaces by $L^p(0,T;Y)$ and $C([0,T];Y)$. The norms of the Lebesgue spaces $L^p(0,T)$ and $L^p(\Omega)$ are denoted, as usual, by $\|\cdot\|_{L^p(0,T)}$ and $\|\cdot\|_{L^p(\Omega)}$ and we write $\|\cdot\|_p$ if the domain of integration is clear from the context. Furthermore, we define the L^p-norm of a vector valued function $u \in L^p(\Omega)^m = L^p(\Omega;\mathbb{R}^m)$ for later conveniences as $\|u\|_{L^p(\Omega;\mathbb{R}^m)} = \|u\|_p = (\sum_{i=1}^m \|u_i\|_p^p)^{1/p}$ and $\|u\|_\infty = \sum_{i=1}^m \|u_i\|_\infty$. As usual, $H^1(\Omega_T)$ denotes the real Hilbert space of $L^2(\Omega_T)$-functions with weak derivatives in $L^2(\Omega_T)$ with the following scalar product and norm for $f, g \in H^1(\Omega_T)$

$$\begin{aligned}
\langle f, g\rangle_{H^1(\Omega_T)} &= \int_{\Omega_T} \left(f(t,x)g(t,x) + \partial_t f(t,x)\,\partial_t g(t,x) + \sum_{i=1}^n \partial_{x_i} f(t,x)\,\partial_{x_i} g(t,x) \right) \mathrm{d}t\,\mathrm{d}x,\\
\|f\|_{H^1(\Omega_T)} &= \|f\|_{H^1} := \langle f, f\rangle_{H^1(\Omega_T)}^{1/2} = \left(\|f\|_2^2 + \|\mathrm{D}f\|_2^2\right)^{1/2},\\
\langle u, v\rangle_{H^1(\Omega_T)} &= \sum_{i=1}^m \langle u_i, v_i\rangle_{H^1(\Omega_T)}, \quad u, v \in H^1(\Omega_T)^m.
\end{aligned}$$

In passing, we introduce for f defined on Ω_T the derivatives with respect to t, x and (t,x)

$$\partial_t f := \dot{f} := \frac{\partial}{\partial t} f, \quad \nabla f := \nabla_x f = (\partial_{x_1}, \ldots, \partial_{x_n}) f, \quad \mathrm{D}f := (\partial_t, \partial_{x_1}, \ldots, \partial_{x_n}) f.$$

For m-dimensional vector valued functions u, we consider for later conveniences ∇u or $\mathrm{D}u$ as mn or $m(n+1)$ dimensional vector. If f is defined on $]0,T[$ we emphasize this by writing $\frac{\mathrm{d}}{\mathrm{d}t}$ instead of ∂_t for the derivative. In all cases it is clear from the context if the derivatives are meant to be in the classical, weak or distributional sense. Furthermore, we denote by $\operatorname{div} f(x) := \sum_{i=1}^n \partial_{x_i} f(x)$ the divergence of f.

Integrals and the dependencies of functions can be abbreviated if the dependencies are clear from the context. As an example, for f defined on Ω_T we write $\int_{\Omega_T} f(t,x)\,\mathrm{d}t\,\mathrm{d}x = \int_{\Omega_T} f\,\mathrm{d}t\,\mathrm{d}x$ and $\int_\Omega f(t,x)\,\mathrm{d}x = \int_\Omega f(t)\,\mathrm{d}x$ and so on. We also use the common notation with dots to emphasize the dependence of variables, for example the notation $f(t) = f(t,\cdot)$ interprets $f : \Omega_T \to \mathbb{R}$ as a function defined on Ω for some fixed $t \in]0,T[$.

We choose the following affine linear structure on the the drift $B[u](t,x) = M(t,x)u(t,x) + c(t,x)$ and the running cost $\mathcal{R}[u](t,x) = \alpha(t,x)u(t,x) + \beta(t,x)$. Furthermore, we assume

$$a_{ij} \in W^{1,\infty}(\Omega_T), \quad \alpha, u \in L^\infty(\Omega_T)^m, \quad c \in L^\infty(\Omega)^n, \quad M \in L^\infty(\Omega; \mathbb{R}^{n\times m}), \quad \beta \in L^\infty(\Omega_T).$$

2. Analysis of the Fokker–Planck problem

This section is devoted to the analysis of solutions to our FP problem (1.3)–(1.4). The results presented in this section are mainly extensions of well-known results for different FP problems [4, 11] and parabolic problems [22] to our case given by (2.3)–(2.4).

We assume a dispersion $\sigma \in W^{1,\infty}(\Omega_T; \mathbb{R}^{n\times n})$, and require that σ is full rank. Consequently, the diffusion matrix (a_{ij}) is coercive in the sense that there exists $\theta > 0$ such that

$$\theta|\xi|^2 \leqslant \sum_{i,j=1}^n a_{ij}(t,x)\xi_i\xi_j, \quad \xi \in \mathbb{R}^n, \text{ f.a.e } (t,x) \in \Omega_T. \tag{2.1}$$

We also assume that the drift $B[u] : \Omega_T \to \mathbb{R}^n$ is an affine linear function of the m-dimensional control u as follows

$$B[u](t,x) := M(t,x)\,u(t,x) + c(t,x), \quad (t,x) \in \Omega_T, \tag{2.2}$$

where $c \in L^\infty(\Omega_T)^n$ and the matrix valued function $M \in L^\infty(\Omega; \mathbb{R}^{n\times m})$ are given. This structure appears in, *e.g.*, [35] in the context of optimal control of crowd motion. In this case, $B[u](t,x) = u(t,x) + c(t,x)$, where c represents a given velocity field and u models a velocity deviation. The form (2.2) also appears in an epidemiological context in modelling the rate of infections, where c and M characterize the dynamics of the disease, and each component of u takes the role of a mitigation measure, *cf.* [17, 18, 25]. In the epidemiological context, also a state dependent diffusion is used to model uncertainty in the dynamics. Further, in [8] ecological reaction-diffusion problems also coupled with FP equations are considered, and the following dispersion coefficient, among others, is discussed:

$$\sigma(x) = A + B\,\cos\left(\frac{2\pi}{\lambda}x\right), \quad \text{with } A > |B|,\ \lambda \in \mathbb{R}.$$

Now, we suppose $u \in L^\infty(\Omega_T)$, and investigate a weak formulation of (1.3)–(1.4). For this purpose, notice that the FP equation can be written in flux form as $\partial_t p = \nabla \cdot F$. Thus, applying Green's formula and using the boundary conditions (1.4) yields for a test function $\psi \in H^1(\Omega)$ the following

$$\begin{aligned}\int_\Omega \partial_t p\,\psi\,\mathrm{d}x &= -\int_\Omega \sum_{i,j=1}^n \partial_{x_j}(a_{ij}p)\partial_{x_i}\psi\,\mathrm{d}x + \int_\Omega p\,B[u]\cdot\nabla\psi\,\mathrm{d}x \\ &\quad + \int_{\partial\Omega} \sum_{i,j=1}^n \partial_{x_j}(a_{ij}\,p)\psi\,\hat{n}_i\,\mathrm{d}S(x) - \int_{\partial\Omega} p\,B[u]\cdot\hat{n}\,\psi\,\mathrm{d}S(x) \\ &= -\int_\Omega \left(\sum_{i,j=1}^n \partial_{x_j}(a_{ij}(t,x)p)\partial_{x_i}\psi - p\,B[u]\cdot\nabla\psi\right)\mathrm{d}x = -\int_\Omega F\cdot\nabla\psi\,\mathrm{d}x, \text{ a.e. on }]0,T[\,.\end{aligned}$$

This result gives rise to the following bilinear flux-operator

$$\begin{aligned}&\mathcal{F}_t : H^1(\Omega)\times H^1(\Omega) \to \mathbb{R} \quad \text{f.a.e. } t \in]0,T[\,, \\ &\mathcal{F}_t(p,\psi) := \int_\Omega \left[\sum_{i,j=1}^n \partial_{x_i}\big(a_{ij}(t,x)p(x)\big)\,\partial_{x_j}\psi(x) - p(x)\,B[u](t,x)\cdot\nabla\psi(x)\right]\mathrm{d}x. \end{aligned}\tag{2.3}$$

The well-definedness of (2.3) is shown in Lemma 2.2 below. Consequently, given an initial state $p_0 = p_{|t=0}$ on Ω, we have the following weak solution concept for (1.3)–(1.4).

Definition 2.1. We call $p \in W^{1,2}(0,T;H^1(\Omega)') \cap L^2(0,T;H^1(\Omega))$ a weak solution to the Fokker–Planck problem with flux-zero boundary conditions and initial state $p_0 \in L^2(\Omega)$ if there exists some null set $N \subset [0,T]$ such that for all $\psi \in H^1(\Omega)$ and all $t \in [0,T]\backslash N$:

$$\langle \dot{p}(t), \psi\rangle_{H'} + \mathcal{F}_t(p(t),\psi) = 0, \qquad p(0) = p_0 \quad \text{a.e. on } \Omega. \tag{2.4}$$

As usual, $H^1(\Omega)'$ denotes the dual space of $H^1(\Omega)$ with pivot space $L^2(\Omega)$,

$$\begin{aligned}&H^1(\Omega)' := \{f : H^1(\Omega) \to \mathbb{R} \,:\, \langle f,\cdot\rangle_{H'} := f(\cdot) \text{ is linear and continuous}\} \\ &\text{and if } f \in L^2(\Omega) \text{ then } \langle f,\cdot\rangle_{H'} = \langle f,\cdot\rangle_{L^2(\Omega)}.\end{aligned}\tag{2.5}$$

In the following, we use the abbreviation $W(0,T) := W^{1,2}(0,T;H^1(\Omega)') \cap L^2(0,T;H^1(\Omega))$ and for later convenience, we define

$$\|p\|_{W(0,T)} := \|p\|_{L^\infty(0,T;L^2(\Omega_T))} + \|p\|_{L^2(0,T;H^1(\Omega))} + \|\dot{p}\|_{L^2(0,T;H^1(\Omega)')}.$$

Furthermore, recall the continuous embedding $W(0,T) \subset C([0,T];L^2(\Omega))$, which gives meaning to the expression $p_{|t=0} \in L^2(\Omega)$. We also remark that the first equation in (2.4) is equivalent to the Bochner-space formulation that is used frequently throughout this paper

$$\dot{p} + \mathcal{F}(p,\cdot) = 0 \quad \text{in } L^2(0,T;H^1(\Omega)') \tag{2.6}$$

and the well-known distributional formulation

$$\partial_t \langle p, \psi\rangle_{H'} + \mathcal{F}(p,\psi) = 0 \quad \text{in } \mathcal{D}'(0,T),\ \psi \in H^1(\Omega). \tag{2.7}$$

In order to show existence of a unique weak solution, we need certain *a priori* bounds. Since the diffusion does depend on x, we need to exploit the coercivity of the diffusion matrix and to this aim, we rewrite the bilinear flux-operator as follows

$$\mathcal{F}(p,\psi) = \int_\Omega \left(\sum_{i,j=1}^n a_{ij}\partial_{x_j} p\, \partial_{x_i}\psi - p\, b\cdot\nabla\psi \right) \mathrm{d}x, \qquad p,\psi \in H^1(\Omega) \tag{2.8}$$

$$\text{with} \quad b_i(t,x) := B[u]_i(t,x) - \sum_{j=1}^n \partial_{x_j} a_{ij}(t,x), \qquad \text{for } (t,x)\in\Omega_T,\ i=1,\dots,n. \tag{2.9}$$

Lemma 2.2. *The flux-operator $\mathcal{F}$ is bounded and weakly coercive, i.e. there exists a null set $N\subset[0,T]$ and $C,\beta,\gamma>0$ such that for all $p,\psi\in H^1(\Omega)$, $t\in[0,T]\backslash N$*

$$\begin{aligned} &\beta\|p\|_{H^1}^2 \leqslant \mathcal{F}_t(p,p) + \gamma\|p\|_{L^2}^2 && \textit{(weak coercivity)},\\ &|\mathcal{F}_t(p,\psi)| \leqslant C\|p\|_{H^1}\|\psi\|_{H^1} && \textit{(boundedness)}. \end{aligned}$$

Proof. We use the ellipticity (2.1) of σ with $\xi := \nabla p$ and obtain f.a.e. $t\in[0,T]$

$$\int_\Omega \theta|\nabla p(x)|^2\,\mathrm{d}x \leqslant \int_\Omega \sum_{i,j=1}^n a_{ij}(t,x)\partial_{x_i}p(x)\,\partial_{x_j}p(x)\,\mathrm{d}x \leqslant \mathcal{F}_t(p,p) + \|b\|_{L^\infty(\Omega_T)}\int_\Omega |p(x)||\nabla p(x)|\,\mathrm{d}x.$$

Next, we use the ε-Young's inequality $ac \leqslant \varepsilon a^2 + c^2/(4\varepsilon)$, which holds for $a,c\in\mathbb{R}$ and $\varepsilon>0$ arbitrary, and choose $a=|\nabla p|$, $c=|p|$ and $\varepsilon = \theta/(2\|b\|_\infty)$. Thus, we conclude

$$\frac{\theta}{2}\int_\Omega |\nabla p(x)|^2\,\mathrm{d}x \leqslant \mathcal{F}_t(p,p) + \|p\|_2^2\|b\|_\infty/(4\varepsilon), \quad \text{f.a.e. } t\in[0,T]. \tag{2.10}$$

Finally, adding $\frac{\theta}{2}\int_\Omega |p(x)|^2\,\mathrm{d}x$ to both sides of (2.10) yields the assertion with constants $\beta := \theta/2$ and $\gamma := \|b\|_\infty^2/(2\theta) + \theta/2$.

In order to show boundedness, let $p,\psi\in H^1(\Omega)$ and obtain

$$\begin{aligned} \mathcal{F}_t(p,\psi) &= \int_\Omega \left(\sum_{i,j=1}^n a_{ij}(t,x)\partial_{x_i}p(x)\,\partial_{x_j}\psi(x) + p(x)\,b(t,x)\cdot\nabla\psi(x) \right)\mathrm{d}x \\ &\leqslant \left(\sum_{i,j=1}^n \|a_{ij}\|_{L^\infty(\Omega_T)} + \sum_{i=1}^n \|b_i\|_{L^\infty(\Omega_T)} \right)\|p\|_{H^1}\|\psi\|_{H^1}. \end{aligned}$$

□

The two properties of the bilinear form $\mathcal{F}$ given in Lemma 2.2 yield existence of a unique solution to (2.4) in the following way.

Theorem 2.3. *For every initial state $p_0 \in L^2(\Omega)$ it holds:*

a) There exists a weak solution $p \in W^{1,2}(0,T;H^1(\Omega)') \cap L^2(0,T;H^1(\Omega))$ of the Fokker–Planck problem with flux-zero boundary conditions and $p(0) = p_0$ in the sense of Definition 2.1.

b) There exists some $C > 0$ only depending on σ, b and Ω_T such that

$$\|p\|_{L^\infty(0,T;L^2(\Omega))} + \|p\|_{L^2(0,T;H^1(\Omega))} + \|\dot{p}\|_{L^2(0,T;H^1(\Omega)')} \leqslant C\|p_0\|_{L^2}$$

c) p is unique in $C([0,T];L^2(\Omega))$.

Proof. The existence of weak solutions can be shown with a standard Galerkin approach. For the convenience of the reader, we prove the necessary *a priori* estimates, even though the results are well-known.

Let p denote a weak solution. Due to the regularity of p, there exists one representative in $C([0,T];L^2(\Omega))$, which is fixed from now on. This yields the following well-known identities, which are used frequently throughout this work, *cf.* [21],

$$p(t) = p(s) + \int_s^t \dot{p}(\tau)\,\mathrm{d}\tau \quad\text{and}\quad \|p(t)\|_2^2 = \|p(s)\|_2^2 + 2\int_s^t \langle \dot{p}(\tau), p(\tau)\rangle_{H'}\,\mathrm{d}\tau, \quad s,t \in [0,T]. \tag{2.11}$$

For a.e. $\tau \in [0,T]$, we can choose $p(\tau) \in H^1(\Omega)$ as test function to obtain with (2.11), Definition 2.1 and the weak coercivity of $\mathcal{F}_t$, the following estimate

$$\|p(t)\|_2^2 = \|p_0\|_2^2 - 2\int_0^t \mathcal{F}_\tau\big(p(\tau), p(\tau)\big)\,\mathrm{d}\tau \leqslant \|p_0\|_2^2 + 2\int_0^t \gamma\|p(\tau)\|_2^2\,\mathrm{d}\tau.$$

Since $t \mapsto \|p(t)\|_{L^2}$ is continuous, we have by Grönwall's inequality

$$\|p(t)\|_2^2 \leqslant e^{2\gamma t}\,\|p_0\|_2^2, \quad t \in [0,T]. \tag{2.12}$$

In order to verify the $L^2(0,T;H^1(\Omega))$-bound we notice with Lemma 2.2, (2.11) and (2.12) that

$$\begin{aligned}\int_0^T \beta\|p(t)\|_{H^1}^2\,\mathrm{d}t &\leqslant \int_0^T \Big(\gamma\|p(t)\|_2^2 - \langle \dot{p}(t), p(t)\rangle_{H'}\Big)\,\mathrm{d}t \\ &\leqslant \int_0^T \gamma e^{2\gamma t}\,\|p_0\|_2^2\,\mathrm{d}t - \frac{1}{2}\big(\|p(T)\|_2^2 - \|p_0\|_2^2\big) \leqslant \frac{1}{2}e^{2\gamma T}\,\|p_0\|_2^2.\end{aligned} \tag{2.13}$$

For a $H^1(\Omega)'$-bound, we only use $\langle \dot{p}(t), \psi\rangle_{H'} = -\mathcal{F}_t(p(t),\psi)$ (f.a.e. t) and the boundedness of $\mathcal{F}_t$ to obtain

$$\|\dot{p}(t)\|_{H^1(\Omega)'} = \sup_{\psi \in H^1(\Omega)} \frac{|\langle \dot{p}(t), \psi\rangle_{H'}|}{\|\psi\|_{H^1}} \leqslant C\|p(t)\|_{H^1}, \quad t \in [0,T]\backslash N. \tag{2.14}$$

Consequently, the $L^2(0,T;H^1(\Omega)')$-bound follows from the $L^2(0,T;H^1(\Omega))$-bound of p and the proof of b) is complete.

In order to verify uniqueness, assume that $p, \tilde{p} \in C([0,T];L^2(\Omega))$ are both weak solutions to the same initial state p_0. Once again with (2.11) and Lemma 2.2 we have

$$\|p(t) - \tilde{p}(t)\|_{L^2}^2 = -2\int_0^t \mathcal{F}_\tau\big(p(\tau) - \tilde{p}(\tau), p(\tau) - \tilde{p}(\tau)\big)\,\mathrm{d}\tau \leqslant C\int_0^t \|p(\tau) - \tilde{p}(\tau)\|_{L^2}^2\,\mathrm{d}\tau.$$

Thus, applying Grönwall's lemma gives the assertion and the proof is complete. ∎

In Lemma 2.2 and Theorem 2.3, we have stated properties that are typical for many parabolic problems. However, in the following corollary, we underline the probabilistic nature of weak solutions to a specific formulation of our Fokker–Planck problem.

Corollary 2.4. *Let $p_0 \in L^2(\Omega)$ be a probability distribution function, i.e.,*

i) $\int_\Omega p_0(x)\,\mathrm{d}x = 1$ *and*
ii) $p_0 \geqslant 0$ *a.e. on* Ω

and let p be the unique weak solution in $C([0,T];L^2(\Omega))$. Then for all $t \in [0,T]$, $p(t) \in L^2(\Omega)$ does also have these properties, in particular,

$$\int_\Omega p(t,x)\,\mathrm{d}x = 1, \quad t \in [0,T].$$

We say that the Fokker–Planck problem with flux-zero boundary conditions is conservative.

Proof. First notice that due to the flux zero boundary condition, the test function appears only as a gradient in the bilinear form $\mathcal{F}$. Hence, the conservation of the total probability follows from the definition of a weak solution if we choose $\psi = 1 \in H^1(\Omega)$ as a test function

$$0 = -\int_s^t \mathcal{F}_\tau(p(\tau),\psi)\,\mathrm{d}\tau = \int_s^t \langle \dot p(\tau),\psi\rangle_{H'}\,\mathrm{d}\tau = \int_\Omega p(t)\,\mathrm{d}x - \int_\Omega p(s)\,\mathrm{d}x, \quad 0 \leqslant s,t \leqslant T. \tag{2.15}$$

But since $\dot p(t)$ is only a $H^1(\Omega)'$-function and since the following argument appears multiple times in this paper we carefully prove the last equal sign. First, recall the continuous embedding

$$C^1([0,T];H^1(\Omega)) \subset W^{1,2}(0,T;H^1(\Omega)') \cap L^2(0,T;H^1(\Omega)),$$

and the fundamental theorem of calculus for Banach space valued functions

$$\varphi(t) - \varphi(s) = \int_s^t \dot\varphi(\tau)\,\mathrm{d}\tau, \quad \text{a.e on } \Omega, \text{ for all } 0 \leqslant s,t \leqslant T,\ \varphi \in C^1([0,T];H^1(\Omega)). \tag{2.16}$$

Hence, the last equal sign in (2.15) follows with a density argument: Let $\left(p_k\right)_{k\in\mathbb{N}} \subset C^1([0,T];H^1(\Omega))$ with $p_k \to p$ in $W^{1,2}(0,T;H^1(\Omega)') \cap L^2(0,T;H^1(\Omega))$. Now by (2.5), Fubini and (2.16) we have

$$\int_s^t \langle \dot p_k(\tau),1\rangle_{H'}\,\mathrm{d}\tau = \int_s^t \langle \dot p_k(\tau),1\rangle_{L^2(\Omega)}\,\mathrm{d}\tau = \int_\Omega \int_s^t \dot p_k(\tau)\,\mathrm{d}\tau\,\mathrm{d}x = \int_\Omega p_k(t)\,\mathrm{d}x - \int_\Omega p_k(s)\,\mathrm{d}x.$$

Taking the limit on both sides proves the conservation of the total probability.

In order to show the non-negativity of p, we consider its negative part

$$p_- := \min\{p,0\} \in L^2(0,T;H^1(0,T)) \cap L^\infty(0,T;L^2(\Omega)).$$

Note that in general p_- does not belong to $W^{1,2}(0,T;H^1(\Omega)')$, nevertheless an integration-by-parts formula still holds and we refer to [40] for a proof. This implies

$$\langle \dot{p_-}(t),p_-(t)\rangle_{H'} = \langle \dot p(t),p_-(t)\rangle_{H'}, \text{ and } \mathcal{F}_t(p(t),p_-(t)) = \mathcal{F}_t(p_-(t),p_-(t)), \quad \text{f.a.e. } t \in]0,T[\,.$$

This yields with $p_-(0)=0$ and the weak coercivity of $\mathcal{F}$ that for every $t\in[0,T]$

$$\frac{1}{2}\|p_-(t)\|_2^2=\int_0^t\langle \dot{p_-}(\tau),p_-(\tau)\rangle_{H'}\,\mathrm{d}\tau=-\int_0^t\mathcal{F}_\tau(p_-(\tau),p_-(\tau))\,\mathrm{d}\tau\leqslant\gamma\int_0^t\|p_-(\tau)\|_2^2\,\mathrm{d}\tau.$$

Now, Grönwall's inequality implies that $\|p_-(t)\|_2^2\leqslant 0$ which in turn provides $p(t)\geqslant 0$ a.e. on Ω. □

Next, we investigate further regularity properties of weak solutions to (2.4). For this purpose, let us recall some continuous Sobolev-embedings, *cf.* [1],

$$H^1(\Omega)\subset\begin{cases}C(\Omega), & & \text{if } n=1,\\ L^\eta(\Omega), & \eta\in[1,\infty[& \text{if } n=2,\\ L^{p^*}(\Omega), & p^*:=\frac{2n}{n-2} & \text{if } n\geqslant 3.\end{cases}\tag{2.17}$$

Corollary 2.5. *(Further regularity of $W(0,T)$-functions)*
Let $n\geqslant 3$ (the dimension of Ω) and $p^*:=\frac{2n}{n-2}\in]2,6]$. Then any function in $W(0,T)$ is also in the following $L^\tau(0,T;L^r(\Omega))$ spaces and we have for all $\alpha\in]0,1[$ the compact embeddings

$$\begin{aligned}&W(0,T)\Subset L^\tau(0,T;L^r(\Omega)), && 1\leqslant\tau<\frac{2}{\alpha},\ 1\leqslant r<\frac{2p^*}{2\alpha-p^*(\alpha-1)}=\frac{2n}{n-2\alpha},\\ &W(0,T)\Subset L^\eta(\Omega_T), && 1\leqslant\eta<4-\frac{4}{p^*}=\frac{4}{n}+2.\end{aligned}$$

If $n\in\{1,2\}$ we even have $W(0,T)\Subset L^\eta(\Omega_T)$ for all $1\leqslant\eta<\infty$.

Proof. The Rellich-Kondrachov compact embedding $H^1(\Omega)\Subset L^{p^*}(\Omega)$ implies the compact embedding $L^2(0,T;H^1(\Omega))\Subset L^2(0,T;L^{p^*}(\Omega))$. Furthermore, for any bounded sequence $(z_k)\subset W(0,T)$, we have for a subsequence

$$z_k\to z\quad\text{in } L^2(0,T;L^{p^*}(\Omega))\quad\text{and}\quad |z_k-z| \text{ is uniformly bounded in } L^\infty(0,T;L^2(\Omega)).$$

With a standard interpolation estimate for Bochner-spaces, we obtain

$$\|z_k-z\|_{\tau,r}\leqslant\|z_k-z\|_{\infty,2}^{1-\alpha}\|z_k-z\|_{2,p^*}^{\alpha}\to 0,$$

where $\frac{1}{\tau}=\frac{1-\alpha}{\infty}+\frac{\alpha}{2}$ and $\frac{1}{r}=\frac{1-\alpha}{2}+\frac{\alpha}{p^*}$. Rearranging both equations to τ and r yields the upper bounds in the first line. The second embedding is the case $\tau=r$. □

3. An inhomogeneous Fokker–Planck problem

In this section, in preparation of our analysis of optimality conditions, we discuss an inhomogeneous FP equation with a right-hand side belonging to the space $L^2(0,T;H^1(\Omega)')$. The presence of a source term leads to the fact that the FP solution is no longer a PDF and Corollary 2.4 does not hold in general.

The main result of this section is the L^∞-estimate given in Theorem 3.2, which is essential for the upcoming analysis of the FP ensemble optimal control problem. Although new, this result is known to be true for similar parabolic equations and we were able to use the available techniques of the proof to our case; see [13, 41].

As in the previous section, we have the general assumptions

$$a_{ij}\in W^{1,\infty}(\Omega_T),\quad u\in L^\infty(\Omega_T)^m,\quad c\in L^\infty(\Omega)^n,\quad M\in L^\infty(\Omega;\mathbb{R}^{n\times m}),$$

and $B[u]$, $\mathcal{F}$, b are given by (2.2), (2.8), (2.9).

Corollary 3.1.
Let $g \in L^2(0,T;H^1(\Omega)')$ and $p_0 \in L^2(\Omega)$. Then there exists a unique weak solution $p \in W(0,T)$ of the inhomogeneous Fokker–Planck problem in the sense of Definition 2.1, such that there exists a null set $N \subset [0,T]$ with

$$\langle \dot{p}(t), \psi \rangle_{H'} + \mathcal{F}_t(p(t), \psi) = \langle g(t), \psi \rangle_{H'}, \quad t \in [0,T] \backslash N,\ \psi \in H^1(\Omega) \tag{3.1}$$

with initial condition $p(0) = p_0$ a.e. on Ω. Additionally, there exists some $C > 0$ only depending on $\Omega_T, \sigma, c, M, \|u\|_\infty$ such that

$$\|p\|^2_{L^\infty(0,T;L^2(\Omega))} + \|p\|^2_{L^2(0,T;H^1(\Omega))} + \|\dot{p}\|^2_{L^2(0,T;H^1(\Omega)')} \leqslant C\big(\|p_0\|^2_{L^2} + \|g\|^2_{L^2(0,T;H^1(\Omega)')}\big). \tag{3.2}$$

Proof. Due to the linearity of the Fokker–Planck equation, the proof can be easily deduced from the proof of Theorem 2.3. □

The following L^∞-estimate is crucial for the second-order analysis and is shown with a De-Giorgi Iteration.

Theorem 3.2. *(L^∞-estimates for the inhomogeneous Fokker–Planck problem)*
Let $n \in \mathbb{N}$, $p_0 \in L^\infty(\Omega)$ and let $p \in W(0,T) \cap C([0,T];L^2(\Omega))$ be the unique weak solution of the inhomogeneous Fokker–Planck problem

$$\langle \dot{p}, \cdot \rangle_{H'} + \mathcal{F}(p, \cdot) = \langle \mathcal{G}, \cdot \rangle_{H'}, \quad \textit{in } L^2(0,T;H^1(\Omega)')$$

with flux-zero boundary conditions and $p(0) = p_0$ a.e. on Ω. Let the source term be of the form

$$\langle \mathcal{G}(t), \psi \rangle_{H'} := \int_\Omega \Big(g_1(t,x)\, \psi(x) + g_2(t,x) \cdot \nabla \psi(x) \Big)\, \mathrm{d}x, \quad t \in [0,T], \quad \psi \in H^1(\Omega), \tag{3.3}$$

where $g_1, g_2 \in L^q(\Omega_T)$ with $q > \frac{4(n+2)}{n+4}$ if $n \geqslant 3$ and $q > 2$ if $n \in \{1,2\}$. Then there exist some $\gamma, C > 0$ depending only on the initial state, the dispersion and the drift, such that

$$\|p(t)\|_\infty \leqslant e^{\gamma t} \|p_0\|_\infty + C\left(\|g_1\|_q + \|g_2\|_q + \|p\|_q\right), \qquad t \in [0,T]. \tag{3.4}$$

Remark: If $n = 3$ we know that $p \in L^\eta(\Omega_T)$ with $\eta < \frac{10}{3}$. On the other hand, $q > \frac{4(n+2)}{n+4} = \frac{20}{7}$ which is smaller than $\frac{10}{3}$ and hence in (3.4) $\|p\|_q \leqslant C\|p_0\|_2$ for $n \in \{1,2,3\}$.

Proof. We fix a pointwise defined representative of $p \in W(0,T)$ that is in $C([0,T];L^2(\Omega))$ and for any $\gamma > 0$, $\lambda > \|p_0\|_\infty$, we define the $C([0,T];L^2(\Omega))$-functions

$$f(t,x) := e^{-\gamma t} p(t,x), \qquad f_\lambda(t,x) := \max\{f(t,x) - \lambda, 0\}.$$

Notice that $f \in W(0,T)$, f_λ is non-negative on Ω_T and positive on the measurable set

$$M_\lambda := \{(t,x) \in \Omega_T \ :\ f(t,x) > \lambda\}$$

and an integration-by-parts formula holds, *cf.* [40]. We remark that the $(n+1)$-dimensional volume of M_λ does not depend on the choice of the pointwise defined representative of p. Furthermore, we can assume that $\operatorname{vol} M_\lambda > 0$ for all $\lambda > \|p_0\|_\infty$, otherwise the assertion is already shown.

STEP 1: For a.e. $t \in [0,T]$, observe that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\|f_\lambda(t)\|_2^2\right) = \langle \dot{f}_\lambda(t), f_\lambda(t)\rangle = \langle \dot{f}(t), f_\lambda(t)\rangle \tag{3.5}$$

$$= -\gamma \int_\Omega f(t) f_\lambda(t)\,\mathrm{d}x - \mathcal{F}_t(f(t), f_\lambda(t)) + \mathcal{G}_t(f_\lambda(t)), \tag{3.6}$$

since p solves $\langle p, \cdot\rangle = -\mathcal{F}_t(p,\cdot)$ in $L^2(0,T;H^1(\Omega)')$ and $f_\lambda(t) \in H^1(\Omega)$. Due to (2.1), we find that a.e. on $[0,T]$

$$-\mathcal{F}(f,f_\lambda) = -\int_\Omega \left(\sum_{i,j=1}^n a_{ij}\partial_{x_i} f_\lambda \partial_{x_j} f_\lambda - fb\cdot\nabla f_\lambda\right)\mathrm{d}x \leqslant -\theta\int_\Omega |\nabla f_\lambda|^2\,\mathrm{d}x + \int_\Omega fb\cdot\nabla f_\lambda\,\mathrm{d}x. \tag{3.7}$$

In the first step, we use the fact that $f_\lambda = 0$ on $\Omega_T\backslash M_\lambda$ and $\nabla f_\lambda = \nabla f$ on M_λ.

Now since $\lambda > \|p_0\|_\infty$, we have $\|f_\lambda(0)\|_2 = 0$. Combining (3.5) and (3.7) and integrating with respect to t yields

$$\begin{aligned}\frac{1}{2}\|f_\lambda(t)\|_2^2 &= -\gamma\int_0^t\int_\Omega f(s,x)f_\lambda(s,x)\,\mathrm{d}s\,\mathrm{d}x - \int_0^t \mathcal{F}_s(f(s),f_\lambda(s))\,\mathrm{d}s + \int_0^t \mathcal{G}_s(f_\lambda(s))\,\mathrm{d}s\\ &\leqslant \int_0^t\left(-\gamma\lambda\int_\Omega f_\lambda\,\mathrm{d}x - \gamma\|f_\lambda\|_2^2 - \theta\|\nabla f_\lambda\|_2^2 + \int_\Omega \left(g_1 f_\lambda + (g_2 + fb)\cdot\nabla f_\lambda\right)\,\mathrm{d}x\right)\mathrm{d}t,\end{aligned} \tag{3.8}$$

where we suppress the arguments of the functions in the last step for the sake of clarity. We use the ε-Young inequality to obtain on M_λ

$$(g_2 + fb)\cdot\nabla f_\lambda \leqslant \frac{4}{\varepsilon}\left(|fb|^2 + |g_2|^2\right) + 2\varepsilon|\nabla f_\lambda|^2, \qquad g_1 f_\lambda \leqslant \frac{4}{\varepsilon}g_1^2 + \varepsilon f_\lambda^2.$$

Since $-\gamma\lambda\int_\Omega f_\lambda(t,x)\,\mathrm{d}x$ is non-positive, we obtain with (3.8) the following inequality

$$\begin{aligned}\frac{1}{2}\|f_\lambda(t)\|_2^2 \leqslant \int_0^t &\left((\varepsilon-\gamma)\|f_\lambda(s)\|_2^2 + (2\varepsilon-\theta)\|\nabla f_\lambda(s)\|_2^2\right)\mathrm{d}s\\ &+ \frac{4}{\varepsilon}\left(\|g_1\|_{L^2(M_\lambda)}^2 + \|g_2\|_{L^2(M_\lambda)}^2 + \|b\|_\infty^2\|f\|_{L^2(M_\lambda)}^2\right).\end{aligned}$$

Next, the choice $\varepsilon = \theta/4$, $\gamma = \theta/2$ results in both $(2\varepsilon-\theta)$ and $(\varepsilon-\gamma)$ being negative; thus we arrive at

$$\|f_\lambda\|_{\infty,2}^2 + \|f_\lambda\|_{2,H^1}^2 \leqslant C\left(\|g_1\|_{L^2(M_\lambda)}^2 + \|g_2\|_{L^2(M_\lambda)}^2 + \|f\|_{L^2(M_\lambda)}^2\right). \tag{3.9}$$

STEP 2: Since $f_\lambda \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega))$, Corollary 2.5 yields $f_\lambda \in L^\eta(\Omega_T)$ with $1 \leqslant \eta < 2 + 4/n$ and we can bound

$$\|f_\lambda\|_\eta \leqslant C\|f_\lambda\|_{\infty,2}^{1-\alpha}\|f_\lambda\|_{2,H^1}^\alpha \leqslant C\left(\|f_\lambda\|_{\infty,2}^2 + \|f_\lambda\|_{2,H^1}^2\right)^{1/2},$$

where we used Young's inequality for the second estimate. If $n = 1$ or $n = 2$, this estimate holds even for all $1 \leqslant \eta < \infty$. Next, we apply Hölder's inequality with the indicator function and exponent $\kappa' := q/2$ with dual κ

to obtain for $i = 1, 2$

$$\int_{M_\lambda} |g_i|^2 \,\mathrm{d}t\,\mathrm{d}x \leqslant (\,\mathrm{vol}\, M_\lambda)^{1/\kappa} \left(\int_{M_\lambda} |g_i|^{2\kappa'} \,\mathrm{d}t\,\mathrm{d}x \right)^{1/\kappa'},$$
$$\int_{M_\lambda} |f|^2 \,\mathrm{d}t\,\mathrm{d}x \leqslant (\,\mathrm{vol}\, M_\lambda)^{1/\kappa} \left(\int_{M_\lambda} |f|^{2\kappa'} \,\mathrm{d}t\,\mathrm{d}x \right)^{1/\kappa'}.$$

This implies with (3.9) and $q = 2\kappa'$ the estimate

$$\|f_\lambda\|_\eta \leqslant C\,(\,\mathrm{vol}\, M_\lambda)^{1/(2\kappa)}\, G, \qquad \text{with } G := \left(\|g_1\|_{L^q(\Omega_T)} + \|g_2\|_{L^q(\Omega_T)} + \|f\|_{L^q(\Omega_T)} \right). \tag{3.10}$$

Step 3: In this step, we bring the previous results together and consider the well-defined, non-increasing function $\varphi(\lambda) := \mathrm{vol}\, M_\lambda$, defined for $\lambda \in [\,\|f_0\|_\infty, \infty[$. Now let $m > \lambda > \|f_0\|_\infty$. Hence, $M_m \subset M_\lambda$ and on M_m, $m - \lambda \leqslant f - \lambda = f_\lambda$. Thus, by (3.10) we obtain

$$\varphi(m)(m-\lambda)^\eta = \int_{M_m} (m-\lambda)^\eta \leqslant \int_{M_\lambda} f_\lambda^\eta \,\mathrm{d}t\,\mathrm{d}x \leqslant C\,(\varphi(\lambda))^{\eta/(2\kappa)}\, G^\eta$$

and therefore $\varphi(m) \leqslant C\,(G/(m-\lambda))^\eta\,(\varphi(\lambda))^{\eta/(2\kappa)}$.

Lastly, we must verify that $\eta/(2\kappa) > 1$ such that we can apply a De-Giorgi Iteration, see Lemma 4.1.1 of [41]. The case $n \in \{1, 2\}$ gives no upper bound on η and hence $q > 2$. If $n \geqslant 3$, we recall that $2\kappa = \frac{\kappa'}{\kappa'-1} = \frac{q}{q-2}$ and $1 \leqslant \eta < 2 + \frac{4}{n}$ and hence we obtain the condition $\frac{(2+4/n)(q-2)}{q} > 1$, *i.e.*, $q > \frac{4(n+2)}{n+4}$. In view of Lemma 4.1.1 in [41], $\mathrm{vol}\, M_m = 0$ for $m \geqslant \|f_0\|_\infty + CG$.

Analogously, we can show that the set $M_\lambda^- := \{(t, x) \in \Omega_T \,:\, f(t, x) < -\lambda\}$ has measure zero for sufficiently large λ by considering $f_\lambda^- := (f + \lambda)_-$ instead of f_λ, which yields the desired lower bound of f. Combining both results, we have shown that $\|f\|_\infty \leqslant \|f_0\|_\infty + C(\|g_1\|_q + \|g_2\|_q + \|f\|_q)$. □

Remark: The assumption $q > \frac{4(n+2)}{n+4}$ for the case $n \geqslant 3$ can be weakened if $g_2 = 0$ and if we find better estimates for the gradient of p. For example, it is well-known hat one obtains the regularity $p \in L^2(0, T; H^k(\Omega))$ with more assumptions on the Fokker–Planck problem. Since we do not need this for the second-order analysis, we will not go into this any further.

Theorem 3.2 immediately gives L^∞-solutions for the Fokker–Planck problem if we consider controls u and initial data p_0 from $L^\infty(\Omega)$ with dimension $n \leq 3$. Even more far-reaching consequences are, as we show in Lemma 4.2, that even the weak L^2-convergence of controls is enough to obtain the convergence of the corresponding solutions in $W(0, T)$.

4. The Fokker–Planck control-to-state map

In this section, we analyse the mapping of a control u to its corresponding state p solving our Fokker–Planck problem. For this analysis we follow [4] to show the well-definedness and differentiability. Moreover, we obtain compactness of the FP control-to-state map for time and space dependent controls. This is done by using the L^∞-estimates given by Theorem 3.2.

In the following, we assume the given initial state $p_0 \in L^\infty(\Omega)$ to be a probability density function with state dimension $n \leq 3$; see Corollary 2.4. Furthermore, let $u^{\min}, u^{\max} : \overline{\Omega_T} \to \mathbb{R}^m$ be measurable and bounded such that $\emptyset \neq U_{\mathrm{ad}}$, U_{ad}^H with

$$\begin{aligned} U_{\mathrm{ad}} &:= \left\{ u \in L^\infty(\Omega_T)^m \,:\, u_i^{\min} \leqslant u_i \leqslant u_i^{\max}, \quad \text{a.e. on } \Omega_T,\ i = 1, \ldots, m \right\}, \\ U_{\mathrm{ad}}^H &:= U_{\mathrm{ad}} \cap H^1(\Omega_T)^m \end{aligned} \tag{4.1}$$

and define

$$U := \left(H^1 \cap L^\infty(\Omega_T)\right)^m,$$
$$\|u\|_U = \|u\|_{H^1(\Omega_T)} + \|u\|_{L^\infty(\Omega_T)}.$$

In order to give meaning to Fréchet-differentiability on the admissible sets, we assume that the interior of U_{ad} (and U_{ad}^H) with respect to $\|\cdot\|_\infty$ (and $\|\cdot\|_U$) is non-empty. Moreover, we add the control vector u to the notation of the bilinear flux and write $\mathcal{F}[u]$ instead of just $\mathcal{F}$.

Definition 4.1.
There exists a unique, non-linear, continuous mapping

$$G : L^\infty(\Omega_T)^m \to W(0,T), \quad u \mapsto G(u), \tag{4.2}$$

such that $G(u)$ represents the weak solution of the Fokker–Planck problem (1.3)–(1.4):

$$\langle \partial_t \left(G(u)\right), \cdot\rangle_{H'} + \mathcal{F}[u](G(u), \cdot) = 0, \qquad \text{in } L^2(0,T;H^1(\Omega)'), \tag{4.3}$$
$$G(u)(0,x) = p_0(x) \qquad \text{f.a.e. } x \in \Omega. \tag{4.4}$$

The operator G maps any admissible control to the associated state and is therefore referred to as the control-to-state operator.

Existence and uniqueness of G follow directly from Theorem 2.3 and the expression $G(u)(0) \in L^2(\Omega)$ is well defined. Next, we discuss further properties of the control-to-state map G, that is, Fréchet-differentiability, Lipschitz-continuity and compactness.

We start with the Fréchet-differentiability on $L^\infty(\Omega_T)^m$ and U. We follow [4] and consider the functional

$$H : W(0,T) \times L^\infty(\Omega_T)^m \to L^2(0,T;H^1(\Omega)') \times L^2(\Omega), \tag{4.5}$$
$$H(p,u) := \big(H_1(p,u), H_2(p(0))\big) := (\dot{p} + \mathcal{F}[u](p,\cdot), p(0) - p_0). \tag{4.6}$$

First, we see that both components of H are arbitrarily often continuously Fréchet-differentiable on $W(0,T) \times L^\infty(\Omega_T)^m$. Now, observe that H was defined such that $H(p,u) = (0,0)$ iff p is a solution of the Fokker–Planck problem with drift u and initial PDF p_0. Hence, $H(G(u),u) = (0,0)$ for all $u \in L^\infty(\Omega_T)^m$. Next, we recall the implicit function theorem on Banach spaces. In order to apply this theorem, we have to show that the mapping

$$W(0,T) \ni z \mapsto D_pH(p,u)(z) = (\dot{z} + \mathcal{F}[u](z,\cdot), z(0)) \in L^2(0,T;H^1(\Omega)') \times L^2(\Omega) \tag{4.7}$$

is an isomorphism. But this follows immediately from Corollary (3.1), specifically, the injectivity follows by the uniqueness and the surjectivity by the existence result.

Hence, the implicit function theorem is applicable for any starting points $(p,u) \in W(0,T) \times L^\infty(\Omega_T)^m$ with $H(p,u) = (0,0)$. Finally, we can deduce that G is continuously Fréchet-differentiable in $u \in L^\infty(\Omega_T)^m$, if we apply this theorem in $(G(u),u)$. This yields a continuously Fréchet-differentiable function $\tilde{G}$ with $H(\tilde{G}(u),u) = (0,0)$ on a open neighbourhood $u \in \tilde{U} \subset L^\infty(\Omega_T)^m$. By uniquness, $\tilde{G} = G$ on $\tilde{U}$, and since u was chosen arbitrarily, we obtain the differentiability of G on $L^\infty(\Omega_T)^m$.

Furthermore, differentiating $H(G(u),u) = 0$ with respect to u gives an implicit formula for $G'(u)$, namely

$$D_pH(G(u),u)G'(u)(v) + D_uH(G(u),u)(v) = 0, \quad u,v \in L^\infty(\Omega_T)^m. \tag{4.8}$$

First, notice that G maps to $W(0,T)$ and hence for $u,v \in L^\infty(\Omega_T)^m$ we have $G'(u)v \in W(0,T)$. Next, let us calculate the Fréchet derivative of H_1 at (p,u) in direction $v \in L^\infty(\Omega_T)^m$. For any test function $\varphi \in W(0,T)$, we have a.e. on $[0,T]$

$$D_u H_1(p,u)(v)(\varphi) = \lim_{\alpha\to 0} \frac{\mathcal{F}[u+\alpha v](p,\varphi) - \mathcal{F}[u](p,\varphi)}{\alpha} = \int_\Omega p\,(Mv)\cdot\nabla\varphi\,\mathrm{d}x. \tag{4.9}$$

Plugging (4.9) and (4.7) into (4.8) implies that $z := G'(u)v$ solves the so-called linearized state equation at $(G(u),u) =: (p,u)$ in direction $v \in L^\infty(\Omega_T)^m$

$$\langle \dot z, \cdot\rangle_{H'} + \mathcal{F}[u](z,\cdot) = \langle f^{\mathrm{lin}}[u,v], \cdot\rangle_{H'}, \quad \text{in } L^2(0,T;H^1(\Omega)'), \quad z(0) = 0 \text{ a.e. on } \Omega, \tag{4.10}$$

where for $\psi \in H^1(\Omega)$, $t \in [0,T]$

$$\langle f_t^{\mathrm{lin}}[u,v], \psi\rangle_{H'} := -\sum_{i=1}^n \sum_{j=1}^m \int_\Omega p(t,x)\, M_{ji}(t,x) v_j(t,x)\, \partial_{x_i}\psi(x)\,\mathrm{d}x.$$

We notice that all of the above can be done analogously if we replace $L^\infty(\Omega_T)^m$ with U.

For the upcoming first- and second-order analysis, it is essential that (4.10) is an inhomogeneous Fokker–Planck problem and the source term $f^{\mathrm{lin}}[u,v]$ takes the form above. Now, let us summarize our previous and some further results with the following lemma:

Lemma 4.2. *The control-to-state map G is of class $C^\infty(L^\infty(\Omega_T)^m)$. Furthermore, it has the following properties:*

a) Its derivative is the solution of the linearized state equation, i.e., $z := G'(u)v \in W(0,T)$ solves (4.10) for $u, v \in L^\infty(\Omega_T)^m$ and it holds that

$$\|z\|_{L^\infty(0,T;L^2(\Omega))} + \|z\|_{L^2(0,T;H^1(\Omega))} + \|\dot z\|_{L^2(0,T;H^1(\Omega)')} \leqslant C\|v\|_{L^2(\Omega_T)}\|G(u)\|_{L^\infty(\Omega_T)}, \tag{4.11}$$

where $C = C(\Omega_T, \sigma, c, M, \|u\|_\infty)$.

b) G is locally Lipschitz-continuous on $L^\infty(\Omega_T)^m$ w.r.t the $W(0,T)$-norm, i.e.,

$$\|G(u) - G(w)\|^2_{W(0,T)} \leqslant C\|G(w)\|^2_{L^\infty(\Omega_T)}\|u-w\|^2_{L^2(\Omega_T)}, \quad u, w \in L^\infty(\Omega_T)^m. \tag{4.12}$$

c) For $u \in L^\infty(\Omega_T)^m$ and for any sequence $(u^k) \subset U_{\mathrm{ad}}$ with $u^k \rightharpoonup u$ in $L^2(\Omega_T)$ it holds that $G(u^k) \to G(u)$ in $W(0,T)$ for a subsequence.

Proof. In order to prove estimate (4.11), we recall that z is a solution of the Fokker–Planck problem with initial state zero and right-hand side $f^{\mathrm{lin}}[u,v]$. Thus, we can apply the estimate (3.2) from Corollary 3.1 and observe for $\psi \in H^1(\Omega)$

$$\|f^{\mathrm{lin}}[u,v](\psi)\|_{L^2(0,T)} \leqslant \left\| \int_\Omega |p(\cdot,x)v(\cdot,x)\cdot\nabla\psi(x)|\,\mathrm{d}x \right\|_{L^2(0,T)} \leqslant \|p\|_\infty\|v\|_2\|\psi\|_{H^1(\Omega)}, \tag{4.13}$$

and hence (4.11) follows. Notice the following: if $v = v(t) \in L^2(0,T)^m$ the problem becomes less difficult since a bound of p in the $L^\infty(0,T;L^2(\Omega))$-norm is sufficient.

The bound for the derivative of G yields the Lipschitz-continuity in the following way: Let $u, w \in L^\infty(\Omega_T)^m$ and define $z := G(u) - G(w) \in W(0,T)$. A straight forward calculation shows that in the sense of (3.1) from Corollary (3.1), z is the weak solution of the Fokker–Planck problem with drift u and right-hand side $f^{\mathrm{lin}}[w, u-w]$, and hence $z = G'(u)(w-u)$.

Similarly, we show the last assertion and set $z_k := G(u) - G(u^k) = G'(u)v^k$, $v^k := u - u^k$ for $k \in \mathbb{N}$. According to (4.11), z_k is uniformly bounded in $W(0,T)$ and therefore there exists some $z \in L^2(0,T;H^1(\Omega_T))$ and $\zeta \in L^2(0,T;H^1(\Omega)')$ such that for a subsequence

$$z_k \rightharpoonup z \quad \text{in } L^2(0,T;H^1(\Omega_T)), \qquad z_k \to z \quad \text{in } L^2(\Omega_T), \qquad \dot{z}_k \rightharpoonup \zeta \quad \text{in } L^2(0,T;H^1(\Omega)').$$

For convenience, we prove that $\dot{z} = \zeta$: let $\phi \in C_c^\infty(\,]0,T[\,)$ and $\psi \in H^1(\Omega)$ and we interpret the $L^2(0,T;H^1(\Omega))$-function z as $L^2(0,T;H^1(\Omega)')$-function. On the one hand, we have by the weak convergence in $L^2(0,T;H^1(\Omega)')$ that (for a subsequence)

$$\int_0^T \phi(t)\langle \dot{z}_k(t), \psi\rangle_{H'} \,\mathrm{d}t \to \int_0^T \phi(t)\langle \zeta(t), \psi\rangle_{H'} \,\mathrm{d}t \quad \text{and} \tag{4.14}$$

$$\int_0^T \dot{\phi}(t)\langle z_k(t), \psi\rangle_{H'} \,\mathrm{d}t \to \int_0^T \dot{\phi}(t)\langle z(t), \psi\rangle_{H'} \,\mathrm{d}t \qquad \text{as } k \to \infty. \tag{4.15}$$

On the other hand, we have for $k \in \mathbb{N}$

$$\int_0^T \phi(t)\langle \dot{z}_k(t), \psi\rangle_{H'} \,\mathrm{d}t = -\Big\langle \int_0^T \dot{\phi}(t) z_k(t)\,\mathrm{d}t, \psi \Big\rangle_{H'} = -\int_0^T \dot{\phi}(t)\langle z_k(t), \psi\rangle_{H'} \,\mathrm{d}t\,; \tag{4.16}$$

the fact that we can interchange the integral and the continuous function $\langle \cdot, \psi\rangle_{H'}$ can be shown straight forwardly by an approximation with simple functions. Since $\psi \in H^1(\Omega)$ was arbitrary, this implies with (4.14) and (4.15) that

$$\int_0^T \phi(t)\zeta(t)\,\mathrm{d}t = -\int_0^T \dot{\phi}(t) z(t)\,\mathrm{d}t, \quad \text{in } H^1(\Omega)'. \tag{4.17}$$

Finally, $\dot{z} = \zeta$ in $L^2(0,T;H^1(\Omega)')$ follows from the fact that (4.17) holds for every $\phi \in C_c^\infty(\,]0,T[\,)$.

Now, we can show that $f^{\mathrm{lin}}[u^k, v^k] \to 0$ in $L^2(0,T;H^1(\Omega)')$, which yields $z_k \to 0$ in $W(0,T)$ according to Corollary 3.1. Recall the fact that for any dual $1 < p, q < \infty$ and reflexiv banach space X we have that $L^p(0,T;X')$ and $L^q(0,T;X)'$ are isometric isomorph. Hence, for $\varphi \in L^2(0,T;H^1(\Omega))$ we have (we omit the (t,x) argument in the second line)

$$\begin{aligned}
\int_0^T f_t^{\mathrm{lin}}[u^k, v^k](\varphi(t))\,\mathrm{d}t &= \int_{\Omega_T} G(u^k)(t,x)\, v^k(t,x) \cdot \nabla\varphi(t,x)\,\mathrm{d}t\,\mathrm{d}x \\
&\leqslant \int_{\Omega_T} |G(u^k) - G(u)| v^k \cdot \nabla\varphi\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega_T} G(u)\, v^k \cdot \nabla\varphi\,\mathrm{d}t\,\mathrm{d}x.
\end{aligned}$$

The first term can be estimated against $C(u^{\min}, u^{\max})\|G(u^k) - G(u)\|^2_{L^2(\Omega_T)}\|\varphi\|^2_{L^2(0,T;H^1(\Omega))}$ and therefore converges to zero as k tends to infinity (for a subsequence). On the other hand, we must exploit the fact that $G(u) \in L^\infty(\Omega_T)$ and hence the mapping

$$L^2(\Omega_T)^m \ni v \mapsto \int_{\Omega_T} G(u)(t,x)\, v(t,x) \cdot \nabla\varphi(t,x)\,\mathrm{d}t\,\mathrm{d}x \;\in \mathbb{R}$$

is linear and continuous. Consequently, due to the weak convergence of (v^k) in $L^2(\Omega_T)^m$, the second term also tends to zero and we have shown that $f^{\mathrm{lin}}[u^k, v^k] \to 0$ in $L^2(0,T;H^1(\Omega))'$. □

5. A FP ensemble optimal control problem

In this section, we consider the following functional

$$J(p,u) := \int_{\Omega_T} \Big(\alpha(t,x)\cdot u(t,x) + \beta(t,x)\Big) p(t,x)\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega} \mathcal{T}(x) p(T,x)\,\mathrm{d}x + \frac{\gamma}{2}\|u\|^2_{H^1(\Omega_T)}, \tag{5.1}$$

where $\gamma \geqslant 0$, $\beta \in L^\infty(\Omega_T)$, $\alpha \in L^\infty(\Omega_T)^m$, $\mathcal{T} \in L^2(\Omega)$ and $\mathcal{R}[u] := \alpha \cdot u + \beta$ on Ω_T. If we assume $\gamma = 0$ then J shall be defined on $W(0,T) \times L^\infty(\Omega_T)^m$, else on $W(0,T) \times U$. In this section, both cases $\gamma = 0$ and $\gamma > 0$ are discussed, whereas we are restricted to the case $\gamma > 0$ for the second-order analysis in Section 7.

Our FP ensemble optimal control problem requires to minimize (5.1) subject to the differential constraint given by the FP problem (1.3)–(1.4). Thus, we introduce the reduced functional

$$\hat{J}(u) := J(G(u), u), \quad u \in \begin{cases} L^\infty(\Omega_T)^m, & \text{if } \gamma = 0 \\ U, & \text{if } \gamma > 0 \end{cases} \tag{5.2}$$

and hence we consider two FP optimal control problems with control constraints given by (4.1) is reformulated as the minimization problem

$$\inf_{u \in U_{\mathrm{ad}}} \hat{J}(u) \ \text{ if } \gamma = 0 \qquad \text{and} \qquad \inf_{u \in U^H_{\mathrm{ad}}} \hat{J}(u) \text{ if } \gamma > 0. \tag{5.3}$$

Theorem 5.1.
Let J have the form (5.1) with $\gamma = 0$ in i) and $\gamma > 0$ in ii). Then both optimal control problems

$$i) \qquad \min_{u \in U_{\mathrm{ad}}} \hat{J}(u)$$
$$ii) \qquad \min_{u \in U^H_{\mathrm{ad}}} \hat{J}(u)$$

admit at least one solution.

Proof. The proof is divided in three steps and we always start with the case $\gamma = 0$.

Step 1: First we need to make sure that $\hat{J}$ is bounded from below. On one hand, we can estimate $G(u)$ in $W(0,T)$ with Theorem 2.3 for all $u \in L^\infty(\Omega_T)^m$

$$\|G(u)\|_{L^\infty(0,T;L^2(\Omega))} + \|G(u)\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t G(u)\|_{L^2(0,T;H^1(\Omega)')} \leqslant C\|p_0\|_{L^2}, \tag{5.4}$$

where $C = C(\Omega_T, \sigma, c, M, \|u\|_\infty)$. On the other hand, $u \in U_{\mathrm{ad}}$ implies the boundedness of $\|u\|_\infty \leqslant C(u^{\min}, u^{\max})$ and since $\alpha_i, \beta \in L^\infty(\Omega_T)$ and $\mathcal{T} \in L^2(\Omega)$ it follows that

$$I := \inf_{u \in U_{\mathrm{ad}}} \hat{J}(u) > -\infty. \tag{5.5}$$

The case $\gamma > 0$ is done analogously.

Step 2: Thus, there exists a minimizing sequences $\left(u^k\right)_{k\in\mathbb{N}} \subset U_{\mathrm{ad}}$ such that $\hat{J}(u^k) \to I$ as $k \to \infty$. Now observe that $\|u^k\|_\infty$ is uniformly bounded by $u^{\min}, u^{\max}$ and that $U_{\mathrm{ad}} \subset L^2 \cap L^\infty(\Omega_T)^m$ is closed w.r.t $\|\cdot\|_{L^2(\Omega_T)}$. This implies the existence of some $u \in U_{\mathrm{ad}}$ such that (for a subsequence) $u^k \rightharpoonup u$ in $L^2(\Omega_T)$. Hence, according to Lemma 4.2 c) the sequence of corresponding states $G(u^k)$ converges to $G(u)$ in $W(0,T)$.

If $\gamma > 0$ the minimizing sequence is in $U_{\mathrm{ad}} \cap H^1(\Omega_T)^m$ and bounded because for $k \in \mathbb{N}$

$$\frac{\gamma}{2}\|u^k\|^2_{H^1(\Omega_T)} \leqslant \left(\|\alpha\|_\infty \|u^k\|_\infty + \|\beta\|_\infty\right)\|p\|_1 + \|\mathcal{T}\|_2 \|p(T)\|_2 + \hat{J}(u^k) \to C + I, \tag{5.6}$$

where $C = C(\Omega_T, \sigma, c, M, \alpha, \beta, \mathcal{T}, p_0)$.

In both cases $\gamma = 0$ and $\gamma > 0$, the weak lower semicontinuity of $\hat{J}$, which is shown in Step 3, implies

$$I \leqslant J(G(u), u) \leqslant \liminf_{k\to\infty} J(G(u^k), u^k) = I \tag{5.7}$$

and the assertion is proven.

Step 3: Recall that

$$\hat{J}(u^k) = \int_{\Omega_T} \Big(\alpha(t,x)\cdot u^k(t,x) + \beta(t,x)\Big) G(u^k)(t,x)\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T}(x) G(u^k)(T,x)\,\mathrm{d}x + \frac{\gamma}{2}\|u^k\|^2_{H^1(\Omega_T)}$$

and that $u^k \rightharpoonup u$ in $L^2(\Omega_T)$ already implies $G(u^k) \to G(u)$ in $W(0,T)$ for a subsequence. For the last term, we can obviously use the weak lower semicontinuity of the H^1-norm, since in the case $\gamma > 0$, the estimate (5.6) implies $u^k \rightharpoonup u$ in $H^1(\Omega_T)$ for a subsequence. Further, $G(u^k) \to G(u)$ in $C([0,T]; L^2(\Omega))$ and $\mathcal{T} \in L^2(\Omega)$ yield the convergence for the second term. Concerning the difficult part of the first term, we add $\pm\alpha \cdot u^k\, G(u)$ to obtain

$$\begin{aligned}\int_{\Omega_T} \left(\alpha\cdot u^k\, G(u^k) - \alpha\cdot u\, G(u)\right)\,\mathrm{d}t\,\mathrm{d}x \leqslant& \|\alpha\|_\infty \|u^k\|_2 \left\|G(u^k) - G(u)\right\|_2 \\ &+ \int_{\Omega_T} \alpha\cdot\left(u^k - u\right) G(u)\,\mathrm{d}t\,\mathrm{d}x \to 0\end{aligned}$$

as k tends to infinity for a subsequence and the weak lower semicontinuity of $\hat{J}$ is proven. □

Once existence of optimal controls are established, we are interested in necessary and sufficient conditions for local optimality. Furthermore, we introduce the corresponding optimality system.

6. Optimality systems and first-order conditions

We say that $\bar{u} \in U_{\mathrm{ad}}$ (or $\bar{u} \in U^H_{\mathrm{ad}}$) is a local minimum of $\hat{J}$ in the normed subspace $Y \subset L^2(\Omega_T)^m$ (or $Y \subset H^1(\Omega_T)^m$), if there exists some $\delta > 0$ such that

$$\hat{J}(\bar{u}) \leqslant \hat{J}(\bar{u} + h), \quad \text{for all } h \in Y \text{ with } \|h\|_Y < \delta. \tag{6.1}$$

The differentiability of the reduced cost functional $\hat{J}$ allows us to introduce the following first-order necessary optimality condition

$$\hat{J}'(\bar{u})(u - \bar{u}) \geqslant 0, \quad u \in Y. \tag{6.2}$$

Now, we introduce the Lagrange functional

$$L(p, u, q) = \hat{J}(u) + \left\langle \partial_t p - \operatorname{div} F, q\right\rangle_{L^2(\Omega_T)},$$

and consider (formally) its stationary points. The first-order necessary conditions for u solution to (5.3) are given by

$$\begin{aligned} \partial_t p(t,x) &= \operatorname{div} F(t,x), && (t,x) \in \Omega_T, \\ p(0,x) &= p_0(x), && x \in \Omega, \\ F \cdot \hat{n} &= 0, && \text{on }]0,T[\, \times \partial\Omega, \end{aligned} \tag{6.3}$$

$$\begin{aligned} \partial_t q(t,x) + L^* q(t,x) &= -\mathcal{R}[u](t,x), && (t,x) \in \Omega_T, \\ q(T,x) &= \mathcal{T}(x), && x \in \Omega, \\ \nabla q \cdot \hat{n} &= 0, && \text{on }]0,T[\, \times \partial\Omega, \end{aligned} \tag{6.4}$$

$$\big\langle (\alpha - M\nabla q)\, p, v - u \big\rangle_{L^2(\Omega_T)} \geqslant 0, \quad v \in L^\infty(\Omega_T)^m, \quad \text{if } \gamma = 0, \tag{6.5}$$

$$\big\langle (\alpha - M\nabla q)\, p, v - u \big\rangle_{L^2(\Omega_T)} + \gamma \langle u, v - u \rangle_{H^1(\Omega_T)} \geqslant 0, \quad v \in U, \quad \text{if } \gamma > 0. \tag{6.6}$$

where the so-called adjoint operator is defined as follows

$$L^* q(t,x) := \sum_{i,j=1}^n a_{ij}(t,x) \partial^2_{x_i x_j} q(t,x) + \sum_{i=1}^n b_i(t,x) \partial_{x_i} q(t,x)$$

and q is the solution to the adjoint problem (6.4).

Next, we analyze the optimality system (6.3), (6.4), (6.5) and (6.3), (6.4), (6.6). We prove existence of sufficiently regular solutions to the adjoint problem (6.4) and show the well-definedness of (6.5) and (6.6). Furthermore, we establish a criterion whether or not we have equality in (6.5), *cf.* Corollary 6.4 b).

Similar optimality systems are considered in [4, 34]. We derive the weak formulation of the adjoint problem straight forwardly: Let q be a solution to (6.4) and let $\psi \in C^1(\Omega)$. Integrating by parts for the diffusion term combined with the Neumann-boundary conditions, we have formally for $t \in [0,T]$

$$\int_\Omega (L^* q(t,x)) \psi(x)\, \mathrm{d}x = -\int_\Omega \nabla q(t,x) \cdot \nabla \psi(x)\, \mathrm{d}x + \int_\Omega \psi(x) b(t,x) \cdot \nabla q(t,x)\, \mathrm{d}x = -\mathcal{F}_t[u](\psi, q(t)).$$

This leads to the following concept of weak solutions to the adjoint problem.

Definition 6.1. For any control $u \in L^\infty(\Omega_T)^m$ we say that $q \in W(0,T)$ is the weak solution of the adjoint system (6.4) if there exists some null set $N \subset [0,T]$ such that for all $t \in [0,T] \backslash N$, $\psi \in H^1(\Omega)$

$$-\langle \dot{q}(t), \psi \rangle_{H'} + \mathcal{F}_t[u](\psi, q(t)) = \langle \mathcal{R}[u](t), \psi \rangle_{L^2(\Omega)}, \tag{6.7}$$

$$q(T) = \mathcal{T}, \quad \text{a.e. on } \Omega. \tag{6.8}$$

Theorem 6.2. *For every $\mathcal{T} \in L^2(\Omega)$, $u, \alpha \in L^\infty(\Omega)^m$, $\beta \in L^\infty(\Omega_T)$ there exists a unique solution $q \in W(0,T)$ of (6.7)–(6.8) and it satisfies the estimate*

$$\|q\|_{W(0,T)} \leqslant C\big(\|\mathcal{T}\|_2 + \|u\|_2 + 1\big) \tag{6.9}$$

with $C = C(\Omega_T, \sigma, c, M, \|u\|_\infty, \alpha, \beta)$.

Proof. This follows similarly to the first section after the time transformation $t \mapsto T-t$; notice that

$$\|\mathcal{R}[u]\|^2_{L^2(0,T;H^1(\Omega)')} \leqslant \int_0^T \|\alpha \cdot u + \beta\|^2_{L^2(\Omega)} \,\mathrm{d}t \leqslant C\left(\|\alpha\|^2_\infty \|u\|^2_2 + \|\beta\|^2_\infty\right).$$

□

This result implies on the one hand the existence of weak solutions of (6.4) and on the other hand we have for $u, v \in L^\infty(\Omega_T)^m$ and $p = G(u)$

$$\left\langle (\alpha - M\nabla q)\, p, v-u \right\rangle_{L^2(\Omega_T)} \leqslant (\|\alpha_\infty + \|d\|_\infty \|\nabla q\|_2)\, \|p\|_\infty \|v-u\|_2 < \infty.$$

Next, we show why the first-order condition (6.2) can be formulated as (6.5) or (6.6), respectively.

First, we derive an important connection between the linearized state equation and the adjoint system of the Fokker–Planck problem. Recall that $z := G'(u)v$ for $u, v \in L^\infty(\Omega_T)^m$ solves

$$\dot z + \mathcal{F}[u](z,\cdot) = f^{\mathrm{lin}}[u,v], \qquad \text{in } L^2(0,T;H^1(\Omega)') \tag{6.10}$$

$$z(0) = 0, \qquad \text{a.e. on } \Omega, \tag{6.11}$$

where the right-hand side of the linearized equation is on $[0,T]$

$$f^{\mathrm{lin}}[u,v](\psi) = -\int_\Omega G(u)(\cdot,x)\ (M(\cdot,x)v(\cdot,x)) \cdot \nabla\psi(x)\,\mathrm{d}x, \quad \psi \in H^1(\Omega).$$

Now, notice that

$$\int_{\Omega_T} \mathcal{R}[u](t,x) z(t,x)\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T}(x) z(T,x)\,\mathrm{d}x = \int_0^T f^{\mathrm{lin}}[u,v](q(t))\,\mathrm{d}t, \tag{6.12}$$

which follows by testing the weak formulations of z and q, *cf.* (6.10) and (6.7), with the $H^1(\Omega)$-functions $q(t)$ and $z(t)$. We remark that, due to the regularity $z, q \in W(0,T)$, it holds that

$$\int_0^T \langle \dot q(t), z(t)\rangle_{H'}\,\mathrm{d}t = -\int_0^T \langle q(t), \dot z(t)\rangle_{H'}\,\mathrm{d}t + z(T)q(T).$$

This relation helps us rewriting the derivative of the reduced cost functional $\hat J$ in (6.2). We have

Lemma 6.3. *For $u \in U$ or $u \in L^\infty(\Omega_T)^m$ we define the vector-valued function*

$$\Phi[u] := (\alpha - M\nabla q)\, G(u), \quad \text{on } \Omega_T, \tag{6.13}$$

where q is the corresponding weak solution of the adjoint problem (6.7), (6.8) with control u. Then the Fréchet-derivative of $\hat J$ at u is given by

$$\hat J'(u)v = \int_{\Omega_T} \Phi[u]\cdot v\,\mathrm{d}t\,\mathrm{d}x + \gamma\langle u, v\rangle_{H^1(\Omega_T)}, \quad v \in U \quad \text{if } \gamma > 0, \tag{6.14}$$

$$\hat J'(u)v = \int_{\Omega_T} \Phi[u]\cdot v\,\mathrm{d}t\,\mathrm{d}x, \qquad v \in L^\infty(\Omega_T)^m \quad \text{if } \gamma = 0. \tag{6.15}$$

Proof. First, recall that

$$\hat{J}(u) = \int_{\Omega_T} \mathcal{R}[u](t,x) G(u)(t,x)\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega} \mathcal{T}(x) G(u)(T,x)\,\mathrm{d}x + \frac{\gamma}{2}\|u\|^2_{H^1(\Omega_T)}$$

and hence for $v \in U$ (if $\gamma > 0$) or $v \in L^\infty(\Omega_T)^m$ (if $\gamma = 0$) we obtain with (6.12) and $z := G'(u)v$

$$\begin{aligned}\hat{J}'(u)v &= \int_{\Omega_T} \alpha \cdot v G(u)\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega_T} \mathcal{R}[u] z\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega} \mathcal{T} z(T)\,\mathrm{d}x + \gamma\langle u, v\rangle_{H^1} \\ &= \int_{\Omega_T} \alpha \cdot v G(u)\,\mathrm{d}t\,\mathrm{d}x + \int_0^T f^{\mathrm{lin}}[u,v](q)\,\mathrm{d}t + \gamma\langle u, v\rangle_{H^1} = \int_{\Omega_T} \Phi[u]\cdot v\,\mathrm{d}t\,\mathrm{d}x + \gamma\langle u, v\rangle_{H^1}.\end{aligned}$$

□

This gives an explicit first-order optimality condition:

Corollary 6.4.

a) Let $\gamma > 0$ and let $\bar{u} \in U^H_{\mathrm{ad}}$ be a local minimum of $\hat{J}$ in U. Then

$$\int_{\Omega_T} \Phi[\bar{u}]\cdot(u-\bar{u})\,\mathrm{d}t\,\mathrm{d}x + \gamma\langle \bar{u}, u-\bar{u}\rangle_{H^1} \geqslant 0, \quad u \in U. \tag{6.16}$$

b) Let $\gamma = 0$ and let $\bar{u} \in U_{\mathrm{ad}}$ be a local minimum of $\hat{J}$ in $L^\infty(\Omega_T)^m$. Then

$$\int_{\Omega_T} \Phi[\bar{u}]\cdot(u-\bar{u})\,\mathrm{d}t\,\mathrm{d}x \geqslant 0, \quad u \in L^\infty(\Omega_T)^m. \tag{6.17}$$

Furthermore, it holds for all $i = 1,\dots,m$ and a.e. $(t,x) \in \Omega_T$ that

$$\begin{cases} \Phi_i[\bar{u}](t,x) > 0 & \Longrightarrow \bar{u}_i(t,x) = u_i^{\min}(t,x), \\ \Phi_i[\bar{u}](t,x) < 0 & \Longrightarrow \bar{u}_i(t,x) = u_i^{\max}(t,x), \\ u_i^{\min}(t,x) < \bar{u}_i(t,x) < u_i^{\max}(t,x) & \Longrightarrow \Phi_i[\bar{u}](t,x) = 0. \end{cases}$$

Proof. Equations (6.16) and (6.17) are immediately obtained by rewriting (6.2) with (6.14) and (6.15), respectively. The other assertion in b) follows with a proof by contradiction by testing (6.17) with proper $u \in L^\infty(\Omega_T)^m$. □

7. Second-order analysis

We start our second-order analysis stating the following necessary optimality condition.

Corollary 7.1. *(Necessary second-order optimality condition for $\gamma = 0$)*
Let $\bar{u} \in U_{\mathrm{ad}}$ be a local minimum in $Y = L^\infty(\Omega_T)$ in the sense of (6.1) for the $\gamma = 0$ case. Then

$$\hat{J}''(\bar{u})v^2 \geqslant 0, \quad v \in C_{\bar{u}}.$$

The cone of feasible directions and the critical cone at $\bar{u} \in U_{\mathrm{ad}}$ are defined as

$$\begin{aligned} S_{\bar{u}} &:= \{\lambda(u-\bar{u}) \,:\, \lambda > 0 \text{ and } u \in U_{\mathrm{ad}}\}, \\ C_{\bar{u}} &:= \overline{S_{\bar{u}}}^{L^2(\Omega_T)} \cap \{v \in L^2(\Omega_T)^m \,:\, \hat{J}'(\bar{u})v = 0\}. \end{aligned}$$

Proof. This assertion follows from an application of Theorem 2.2 in [19]; the regularity condition

$$C_{\bar u} = \overline{\{v \in S_{\bar u} \ : \ \hat J'(\bar u)v = 0\}}^{L^2(\Omega_T)}$$

can be shown with standard techniques, *cf.* Theorem 6.6 of [4]. ☐

Next, we prove a sufficient condition for a control $\bar u$ to be locally optimal, *cf.* Theorem 7.2. We only consider the case $\gamma > 0$ and controls in U^H_{ad}.

Therefore, we exploit the following theorem by Casas and Tröltzsch [19]; slightly adapted to our case. In passing, let us introduce $\mathcal{L}(H^1(\Omega_T)^m)$ and $\mathcal{B}(H^1(\Omega_T)^m)$, which denote the space of (bi)linear, continuous functions from $H^1(\Omega_T)^m$ or $H^1(\Omega_T)^m \times H^1(\Omega_T)^m$ to $\mathbb{R}$.

Theorem 7.2. *Let $\bar u \in U^H_{\text{ad}}$. Assume that the reduced cost functional $\hat J : U \to \mathbb{R}$ satisfies the following assumptions:*

(A1) $\hat J$ is of class C^2 in U^H_{ad}. Moreover, for every $u \in U^H_{\text{ad}}$ there exist continuous extensions

$$\hat J'(u) \in \mathcal{L}\left(H^1(\Omega_T)^m\right), \quad \hat J''(u) \in \mathcal{B}\left(H^1(\Omega_T)^m\right). \tag{A1}$$

(A2) There exists $\Lambda > 0$ such that for all $(u^k) \subset U^H_{\text{ad}}$ and $(v^k) \subset U$ with $u^k \to \bar u$ in $H^1(\Omega_T)$ and $v^k \rightharpoonup v$ in $H^1(\Omega_T)$:

$$\hat J'(\bar u)v = \lim_{k\to\infty} \hat J'(u^k)v^k, \tag{A2.1}$$

$$\hat J''(\bar u)v^2 \leqslant \liminf_{k\to\infty} \hat J''(u^k)(v^k)^2. \tag{A2.2}$$

$$\text{and if } v = 0, \text{ then} \quad \Lambda \liminf_{k\to\infty} \|v^k\|^2_{H^1} \leqslant \liminf_{k\to\infty} \hat J''(u^k)(v^k)^2. \tag{A2.3}$$

(A3) $\bar u$ satisfies the first- and second-order necessary conditions

$$\hat J'(\bar u)(u - \bar u) \geqslant 0, \quad u \in U^H_{\text{ad}}, \tag{A3.1}$$

$$\hat J''(\bar u)v^2 > 0, \qquad v \in C_{\bar u}\backslash\{0\}. \tag{A3.2}$$

The sets are defined as follows

$$\begin{aligned}
S_{\bar u} &:= \{\lambda(u - \bar u) \ : \ \lambda > 0 \text{ and } u \in U^H_{\text{ad}}\}, && \textit{(cone of feasible directions)}\\
C_{\bar u} &:= \overline{S_{\bar u}}^{H^1} \cap \{v \in H^1(\Omega_T)^m \ : \ \hat J'(\bar u)v = 0\}, && \textit{(critical cone)}\\
D_{\bar u} &:= \{v \in S_{\bar u} \ : \ \hat J'(\bar u)v = 0\},
\end{aligned}$$

where $\overline{S_{\bar u}}^{H^1}$ denotes the closure of $S_{\bar u}$ with respect to the norm $\|\cdot\|_{H^1(\Omega_T)}$.

Then we have local H^1-optimality at $\bar u$ in the sense that there exists $\varepsilon > 0$ and $\delta > 0$ such that

$$\hat J(\bar u) + \frac{\delta}{2}\|\bar u - u\|^2_{H^1(\Omega_T)} \leqslant \hat J(u), \quad u \in B_\varepsilon(\bar u; H^1(\Omega_T)). \tag{7.1}$$

Proof. We prove by contradiction and assume that (7.1) does not hold. Hence, we find a sequence $(u^k) \subset U^H_{\text{ad}}$ with

$$0 \neq \|\bar u - u^k\|_{H^1} < \frac{1}{k} \quad \text{and} \quad \hat J(\bar u) + \frac{1}{2k}\|\bar u - u^k\|^2_{H^1} > \hat J(u^k), \quad k \in \mathbb{N}.$$

Moreover, we define $v^k := (u^k - \bar{u})/\|u^k - \bar{u}\|_{H^1}$ for $k \in \mathbb{N}$ and thus $\|v^k\|_{H^1} = 1$. Consequently, there exists some $v \in H^1(\Omega_T)^m$ such that, for a subsequence,

$$v^k \rightharpoonup v, \quad \text{in } H^1(\Omega_T)^m, \qquad v^k \to v, \quad \text{in } L^q(\Omega_T)^m, \quad 1 \leqslant q < \frac{2(n+1)}{n-1}.$$

Now if we can show with the help of (A1)–(A3) that $v = 0$, we would obtain the contradiction using assumption (A2.3). To this end, we show that $v \in C_{\bar{u}}$ and $\hat{J}''(\bar{u})v^2 = 0$ which implies $v = 0$ due to (A3.2).

$\underline{v \in C_{\bar{u}}}$: Since (v^k) is a subset of the convex set $S_{\bar{u}}$, we invoke Mazur's Lemma and find a convex combinations that converges strongly to v in $H^1(\Omega_T)^m$. Thus, the closedness of $\overline{S_{\bar{u}}}^{H^1}$ implies $v \in \overline{S_{\bar{u}}}^{H^1}$. In order to show $\hat{J}'(\bar{u})v = 0$, we exploit the necessary first-order condition (A3.1) in the first line and use the mean value theorem with proper coefficients $(\theta_k)_{k\in\mathbb{N}} \subset [0,1]$ in the second line to obtain

$$\begin{aligned} \hat{J}'(\bar{u})v^k &= \frac{1}{\|u^k - \bar{u}\|_{H^1}} \hat{J}'(\bar{u})(u^k - \bar{u}) \geqslant 0, \\ \hat{J}'(\bar{u} + \theta_k(u^k - \bar{u}))v^k &= \frac{\hat{J}(u^k) - \hat{J}(\bar{u})}{\|u^k - \bar{u}\|_{H^1}} < \frac{\|u^k - \bar{u}\|_{H^1}}{2k} \to 0. \end{aligned}$$

Now thanks to Assumption (A2.1), both terms on the left hand side converge to $\hat{J}'(\bar{u})v$ as k tends to infinity, which proves the assertion $\hat{J}'(\bar{u})v = 0$.

$\underline{\hat{J}''(\bar{u})v^2 = 0}$: Once again, we can find $(\theta_k) \subset [0,1]$ such that by a Taylor expansion in $\bar{u}$,

$$\hat{J}(u^k) = \hat{J}(\bar{u} + u^k - \bar{u}) = \hat{J}(\bar{u}) + \hat{J}'(\bar{u})(u^k - \bar{u}) + \frac{1}{2}\hat{J}''(\bar{u} + \theta_k(u^k - \bar{u}))(u^k - \bar{u})^2$$

and hence after dividing by $\|u^k - \bar{u}\|^2_{H^1}$, we obtain

$$\frac{1}{2}\hat{J}''(\bar{u} + \theta_k(u^k - \bar{u}))(v^k)^2 = \frac{\hat{J}(u^k) - \hat{J}(\bar{u})}{\|u^k - \bar{u}\|^2_{H^1}} - \frac{\hat{J}'(\bar{u})v^k}{\|u^k - \bar{u}\|_{H^1}} < \frac{1}{k}.$$

Now, we apply the strong Assumptions (A2.2)

$$0 \leqslant \hat{J}''(\bar{u})v^2 \leqslant \liminf_{k\to\infty} \hat{J}''(\bar{u} + \theta_k(u^k - \bar{u}))(v^k)^2 = 0$$

in order to confirm $v = 0$, and finally (A2.3) gives the desired contradiction

$$0 < \Lambda = \Lambda \liminf_{k\to\infty} \|v^k\|^2_{H^1} \leqslant \liminf_{k\to\infty} \hat{J}''(u^k)(v^k)^2 = 0.$$

□

Let us explain why we are restricted to $\gamma > 0$ and controls from $H^1(\Omega_T)$. On the one hand, we need some L^2-norm on the controls in $\hat{J}$ such that $\hat{J}''$ satisfies assumption (A2.3); notice that this is exactly the assumption which gives the desired contradiction in the proof of Theorem 7.2. On the other hand, replacing $H^1(\Omega_T)$ with $L^2(\Omega_T)$ leads to $v^k \rightharpoonup v$ only in L^2 and we are - in general - no longer able to verify (A2.2). This is due to the fact that v appears on the right hand side of the differential equations which implicitly yield the derivatives of the control-to-state map and certain integrability of it is necessary for L^∞-estimates, see Theorem 7.3.

Now, our aim is to show that our reduced cost functional $\hat{J}$, given by (5.2) for the case $\gamma > 0$, does fulfil the assumptions of Theorem 7.2.

We start by computing the second derivative of the control-to-state map $G : U \to W(0,T)$. Recall that $z = G'(u)v$ solves the linearised state equation (6.10)–(6.11). Thus, differentiating this equation with respect to u gives the following implicit equation for the second derivative of G for $u, v_1, v_2 \in U$

$$\partial_t \left(G''(u)(v_1,v_2)\right) + \mathcal{F}[u]\left(G''(u)(v_1,v_2),\cdot\right) = -\int_{\Omega_T} \left(G'(u)v_1(Mv_2) + G'(u)v_2(Mv_1)\right)\cdot\nabla(\cdot)\,\mathrm{d}t\,\mathrm{d}x. \tag{7.2}$$

Hence, $w := G''(u)(v_1,v_2) \in W(0,T)$ with $z_1 := G'(u)v_1$ and $z_2 := G'(u)v_2$ satisfies the inhomogeneous Fokker–Planck problem

$$\dot{w} + \mathcal{F}[u](w,\cdot) = f^{\text{quad}}_{(u,v_1,v_2)} \quad \text{ in } L^2(0,T;H^1(\Omega))', \tag{7.3}$$

$$w(0) = 0, \quad \text{ in } L^2(\Omega), \tag{7.4}$$

where the right-hand side is defined as follows

$$f^{\text{quad}}_{(u,v_1,v_2)}(\varphi) := -\int_{\Omega_T} (z_1 M v_2 + z_2 M v_1)\cdot\nabla\varphi\,\mathrm{d}t\,\mathrm{d}x, \qquad \varphi \in L^2(0,T;H^1(\Omega)).$$

If $v = v_1 = v_2$, we simply write $f^{\text{quad}}[u,v]$.

Theorem 7.3. *Let $u_1, u_2 \in U^H_{\text{ad}}$ and $v_1, v_2 \in U$ and denote $p_i := G(u_i)$, $z_i := G'(u_i)v_i$, $w_i := G''(u_i)v_i^2$ for $i = 1,2$. Then $z_i, w_i \in L^\infty(\Omega_T)$ and we have*

$$\begin{aligned}
\|z_i\|_\infty + \|w_i\|_\infty &\leqslant C\|v_i\|_q,\\
\|p_1 - p_2\|_\infty &\leqslant C\|M\|_\infty\|u_1-u_2\|_q\|p_1\|_\infty \leqslant C\|u_1-u_2\|_q,\\
\|z_1 - z_2\|_{W(0,T)} &\leqslant \|M\|_\infty\left(\|p_1\|_\infty\|v_1-v_2\|_2 + \|v_2\|_2\|p_1-p_2\|_\infty + \|z_2\|_\infty\|u_1-u_2\|_2\right)\\
&\leqslant C\left(\|v_1-v_2\|_2 + \|v_2\|_2\|u_1-u_2\|_q + \|v_1\|_q\|u_1-u_2\|_2\right),\\
\|z_1 - z_2\|_\infty &\leqslant C\|M\|_\infty\left(\|p_1\|_\infty\|v_1-v_2\|_q + \|v_2\|_q\|p_1-p_2\|_\infty + \|z_2\|_\infty\|u_1-u_2\|_q\right)\\
&\leqslant C\left(\|v_1-v_2\|_q + \|v_2\|_q\|u_1-u_2\|_q + \|v_1\|_q\|u_1-u_2\|_q\right),\\
\|w_1 - w_2\|_{W(0,T)} &\leqslant C\|M\|_\infty\left(\|z_1\|_\infty\|v_1-v_2\|_2 + \|v_2\|_q\|z_1-z_2\|_\infty + \|w_2\|_\infty\|u_1-u_2\|_2\right)\\
\|w_1 - w_2\|_\infty &\leqslant C\|M\|_\infty\left(\|z_1\|_\infty\|v_1-v_2\|_q + \|v_2\|_q\|z_1-z_2\|_\infty + \|w_2\|_\infty\|u_1-u_2\|_q\right)
\end{aligned}$$

for some constant $C = C(\Omega_T, \sigma, c, M, \|p_0\|_\infty, u^{\min}, u^{\max}) > 0$ and exponent $q > \frac{4(n+2)}{n+4} = \frac{20}{7}$ if $n = 3$ and $q > 2$ if $n \in \{1,2\}$.

Notice that $n \leq 3$ implies $\|\cdot\|_q \leqslant C\|\cdot\|_{H^1(\Omega_T)}$ in Theorem 7.3.

Proof. The first assertion can be verified with Theorem 3.2 as follows: In view of (3.3), observe that the right-hand sides $f^{\text{lin}}[u,v]$ and $f^{\text{quad}}[u,v]$ have the correct shape if we define $g_1 = 0$ and $g_2 := -p_i\,M\,v_i$ for z_i and $g_2 := -2z_i\,M v_i$ for w_i. Hence, it holds for any $q > \frac{4(n+2)}{n+4}$ if $n = 3$ and $q > 2$ if $n \in \{1,2\}$ that

$$\begin{aligned}
&\|z_i\|_\infty \leqslant \|g_2\|_q \leqslant \|M\|_\infty\|p\|_\infty\|v_i\|_q \leqslant C\|v_i\|_q,\\
&\|w_i\|_\infty \leqslant \|M\|_\infty\|z_i\|_\infty\|v_i\|_q \leqslant C\|v_i\|_q, \qquad i \in \{1,2\}.
\end{aligned}$$

With the L^∞-estimate for z_1 at hand, we immediately obtain the estimate for $\|p_1 - p_2\|_\infty$ by choosing $v_1 = u_1 - u_2$; recall that $z_1 = G'(u_1)v_1 = p_1 - p_2$, *cf.* Lemma 4.2.

Next, we show the third estimate and define $\delta z := z_1 - z_2$. Similarly to the proof of Lemma 4.2, we obtain that $\delta z \in W(0,T)$ solves again a Fokker–Planck problem with drift u_1 and some right-hand side f^{lin}_δ in

$L^2(0,T;H^1(\Omega)')$. For any $\psi \in H^1(\Omega)$ and a.e. on $[0,T]$, we find that

$$\begin{aligned}\langle \partial_t(\delta z), \psi\rangle_{H'} + \mathcal{F}[u_1](\delta z, \psi) &= -\langle z_2, \psi\rangle_{H'} - \mathcal{F}[u_1](z_2,\psi) + f^{\mathrm{lin}}[u_1,v_1](\psi)\\ &= -f^{\mathrm{lin}}[u_2,v_2](\psi) + \int_\Omega z_2 M(u_1-u_2)\cdot\nabla\psi \,\mathrm{d}x + f^{\mathrm{lin}}[u_1,v_1](\psi)\\ &=: f_\delta^{\mathrm{lin}}(\psi).\end{aligned}$$

According to Corollary 3.1, we need to bound f_δ^{lin} in $L^2(0,T;H^1(\Omega)')$ for the $W(0,T)$-estimate. Almost everywhere on $[0,T]$ we obtain the following estimate

$$\begin{aligned}|f_\delta^{\mathrm{lin}}(\psi)| &\leqslant \int_\Omega |(p_1Mv_1 - p_2Mv_2)\cdot\nabla\psi| \,\mathrm{d}x + \int_\Omega |z_2M(u_1-u_2)\cdot\nabla\psi| \,\mathrm{d}x\\ &\leqslant \|M\|_\infty\Big(\|p_1\|_\infty\|v_1-v_2\|_2 + \|v_2\|_2\|p_1-p_2\|_\infty + \|z_2\|_\infty\|u_1-u_2\|_2\Big)\|\psi\|_{H^1}.\end{aligned}$$

The L^∞-estimate is obtained with Theorem 3.2 if we bound the g_2 term of the right-hand side, *i.e.* $f_\delta^{\mathrm{lin}}(\psi) = \int_{\Omega_T} g_2\cdot\nabla\psi$, in the $L^q(\Omega_T)$-norm as follows

$$\begin{aligned}\|g_2\|_q &= \|p_1Mv_2 - p_1Mv_1 + z_2\,M(u_1-u_2)\|_q\\ &\leqslant \|M\|_\infty\left(\|p_1\|_\infty\|v_1-v_2\|_q + \|v_2\|_q\|p_1-p_2\|_\infty + \|z_2\|_\infty\|u_1-u_2\|_q\right).\end{aligned}$$

The same procedure can be done with $\delta w := w_1 - w_2 \in W(0,T)$ in order to obtain the Lipschitz estimates for G''. For any $\psi\in H^1(\Omega)$ and a.e. on $[0,T]$, we find that

$$\begin{aligned}&\langle \partial_t(\delta w),\psi\rangle_{H'} + \mathcal{F}[u_1](\delta w,\psi)\\ &\qquad = -f^{\mathrm{quad}}[u_2,v_2](\psi) + \int_\Omega w_2M(u_1-u_2)\cdot\nabla\psi\,\mathrm{d}x + f^{\mathrm{quad}}[u_1,v_1](\psi) =: f_\delta^{\mathrm{quad}}(\psi),\end{aligned}$$

where $f_\delta^{\mathrm{quad}} \in L^2(0,T;H^1(\Omega)')$ can be bound analogously to f_δ^{lin} from above. □

With the derivatives of the control-to-state map G at hand, we are ready to compute the second derivative of the reduced cost functional $\hat{J}$. Recall that

$$\hat{J}'(u)v = \int_{\Omega_T} \alpha\cdot vp\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega_T} R[u]z\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T}z(T)\,\mathrm{d}x + \gamma\langle u,v\rangle_{H^1(\Omega_T)}$$

and hence for $u, v_1, v_2 \in U$

$$\hat{J}''(u)(v_1,v_2) = \int_{\Omega_T}\alpha\cdot(v_1z_2+v_2z_1)\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega_T} R[u]w\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega Vw(T)\,\mathrm{d}x + \gamma\|v\|_{H^1}^2 \tag{7.5}$$

We summarize our results and extend the bilinear form $J''(u)$ to U.

Lemma 7.4. *Let $u\in U_{\mathrm{ad}}^H$ and $v\in U$ and denote by $q\in W(0,T)$ the corresponding solution of the adjoint equation. Furthermore, define $p := G(u)$, $z := G'(u)v$ and $w := G''(u)v^2$. Then it holds*

$$\begin{aligned}\hat{J}'(u)v &= \int_{\Omega_T}\alpha\cdot vp\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega_T}\mathcal{R}[u]z\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T}z(T)\,\mathrm{d}x + \gamma\langle u,v\rangle_{H^1(\Omega_T)}\\ &= \int_{\Omega_T}\Phi[u]\cdot v\,\mathrm{d}t\,\mathrm{d}x + \gamma\langle u,v\rangle_{H^1},\end{aligned}$$

$$\begin{aligned}\hat{J}''(u)v^2 &= \int_{\Omega_T} 2\alpha \cdot vz \,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega_T} \mathcal{R}[u]w \,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega} \mathcal{T}w(T)\,\mathrm{d}x + \gamma\|v\|_{H^1}^2 \\ &= \int_{\Omega_T} 2\alpha \cdot vz \,\mathrm{d}t\,\mathrm{d}x - \int_{\Omega_T} 2zMv \cdot \nabla q \,\mathrm{d}t\,\mathrm{d}x + \gamma\|v\|_{H^1}^2.\end{aligned}$$

Proof. Both equations for the first derivative have already been proven in the previous sections. The second formula for $\hat{J}''$ can be shown similarly: Let $u \in U_{\mathrm{ad}}^H$ and $v_1, v_2 \in U$ and $w := G''(u)(v_1, v_2), z_1 := G'(u)v_1, z_2 := G'(u)v_2$. Now, we test the adjoint equation with $w(t)$ and (7.3) with $q(t)$ to obtain a.e. on $[0, T]$

$$\begin{aligned}-\langle \dot{q}, w\rangle_{H'} + \mathcal{F}[u]\,(w, q) &= \int_\Omega R[u]w\,\mathrm{d}x \\ \text{and}\ \ \langle \dot{w}, q\rangle_{H'} + \mathcal{F}[u]\,(w, q) &= -\int_\Omega (z_1 M v_2 + z_2 M v_1)\cdot \nabla q\,\mathrm{d}x.\end{aligned}$$

Using the terminal condition of the adjoint equation $\mathcal{T} = q(T)$, we arrive at

$$\begin{aligned}&\int_{\Omega_T} \mathcal{R}[u]w\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T}w(T)\,\mathrm{d}x = \int_0^T f^{\mathrm{quad}}[u,v](q(t))\,\mathrm{d}t; \quad \text{compare this to} \\ &\int_{\Omega_T} \mathcal{R}[u]z\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T}z(T)\,\mathrm{d}x = \int_0^T f^{\mathrm{lin}}[u,v](q(t))\,\mathrm{d}t\end{aligned}$$

and thus

$$\hat{J}''(u)(v_1, v_2) = \int_{\Omega_T} \alpha\cdot(v_1 z_2 + v_2 z_1)\,\mathrm{d}t\,\mathrm{d}x - \int_{\Omega_T} (z_2 M v_1 + z_1 M v_2)\cdot\nabla q\,\mathrm{d}t\,\mathrm{d}x + \gamma\|v\|_{H^1}^2. \tag{7.6}$$

This immediately implies the boundedness for $v_1, v_2 \in H^1(\Omega_T)^m$; recall that $z_1, z_2 \in L^\infty(\Omega_T)$ and $\nabla q \in L^2(\Omega_T)$. $\square$

The previous Lemma immediately implies that our reduced cost functional $\hat{J}$ fulfils assumption (A1) from Theorem 7.2. Next, let us verify the second assumption (A2).

Lemma 7.5. *The reduced cost functional fulfils assumption (A2).*

Proof. Let $u \in U_{\mathrm{ad}}^H$, $v \in H^1(\Omega_T)^m$ and let $(u^k) \subset U_{\mathrm{ad}}^H$, $(v^k) \subset U$ such that $u^k \to u$ and $v^k \rightharpoonup v$ in $H^1(\Omega_T)^m$. Furthermore, let us define for $k \in \mathbb{N}$

$$p_k := G(u^k), \quad z_k := G'(u^k)v^k, \quad w_k := G''(u^k)(v^k)^2,$$

and p, z, w as usual. Now, Theorem 7.3 implies the convergences (for one subsequence)

$$p_k \to p, \quad z_k \to z, \quad w_k \to w, \quad \text{in } L^\infty(\Omega_T) \text{ and } W(0,T); \tag{7.7}$$

we only demonstrate the uniform convergence of (w_k). First notice that the weak convergence of (v^k) implies $v^k \to v$ in $L^q(\Omega_T)$ for $1 \leqslant q < \frac{2(n+1)}{(n+1)-2} = 4$ if $n = 3$ and for every $1 \leqslant q < \infty$ if $n \in \{1, 2\}$. Thus, Theorem 7.3 yields for $\frac{20}{7} < q < 4$

$$\|z_k\|_\infty + \|w_k\|_\infty \leqslant C\|v^k\|_q \leqslant C\|v\|_q, \quad k \in \mathbb{N}$$

with $C = C(\Omega_T, \sigma, c, M, \|p_0\|_\infty, u^{\min}, u^{\max}) > 0$ and as k tends to infinity we obtain

$$\|w - w_k\|_\infty \leqslant C\|M\|_\infty \Big(\|z\|_\infty \|v - v^k\|_q + \|v^k\|_q \|z - z_k\|_\infty + \|w_k\|_\infty \|u - u^k\|_q \Big) \to 0$$

as desired.

Next, we can prove the following limit

$$\hat{J}'(u^k)v^k = \int_{\Omega_T} \alpha \cdot v^k p_k \,\mathrm{d}t \,\mathrm{d}x + \int_{\Omega_T} \mathcal{R}[u^k] z_k \,\mathrm{d}t \,\mathrm{d}x + \int_\Omega \mathcal{T} z_k(T) \,\mathrm{d}x + \gamma \langle u^k, v^k \rangle_{H^1(\Omega_T)} \to \hat{J}'(u)v$$

as k tends to infinity for a subsequence. The first term converges after inserting $\alpha \cdot v^k p$ and using the $L^\infty(\Omega_T)$-convergence of (p_k). We proceed similarly with the second term. Since $\mathcal{T} \in L^2(\Omega)$, the $L^2(\Omega)$-convergence $z_k(T) \to z(T)$ is sufficient for the third term and the fourth term is standard.

Similarly, we obtain the lower semicontinuity of the second derivative

$$\hat{J}''(u)v^2 \leqslant \liminf_{k\to\infty} \hat{J}''(u^k)(v^k)^2 = \int_{\Omega_T} 2\alpha \cdot v^k z_k \,\mathrm{d}t \,\mathrm{d}x + \int_{\Omega_T} \mathcal{R}[u^k] w_k \,\mathrm{d}t \,\mathrm{d}x + \int_\Omega \mathcal{T} w_k(T) \,\mathrm{d}x + \gamma \|v^k\|_{H^1}^2.$$

The first three terms are dealt with analogously to the above and for the fourth term we use the weak lower semicontinuity of the L^2-norm $\|v\|_{H^1} \leqslant \liminf_{k\to\infty} \|v^k\|_{H^1}$. □

We can conclude this section with the following second-order necessary and sufficient optimality conditions for the Fokker–Planck control problem.

Corollary 7.6. *(Sufficient second-order optimality conditions)*
Let $n \in \{1, 2, 3\}$ and let assumption (A3) from Theorem 7.2 hold for $\bar{u} \in U_{\mathrm{ad}}^H$. Then $\bar{u}$ is a local minimum of $\hat{J}$ in $H^1(\Omega_T)^m$ in the sense of (6.1).

Proof. This is a direct consequence of Theorem 7.2 where Lemma 7.4 and 7.5 guarantee that assumptions (A1) and (A2) are fulfilled. □

8. Extension to Brockett's ensemble control problems

In this section, we discuss optimality conditions in the case $\gamma = 0$ and averaged H^1-costs of the control. Furthermore, we analyze the case of approximate closed-loop controls. We remark that our results in this section apply to the FP ensemble control problems considered in [10, 14–16, 34–36].

In Brockett's framework, an ensemble cost functional includes expected values of the cost of the control. In particular, we consider the following functional

$$J_1(p, u) = \int_{\Omega_T} \Big(\mathcal{R}[u](t, x) + \frac{\gamma_1}{2} |u(t, x)|^2 + \frac{\gamma_2}{2} |\mathrm{D}u(t, x)|^2 \Big) p(t, x) \,\mathrm{d}t \,\mathrm{d}x + \int_\Omega \mathcal{T}(x)\, p(T, x) \,\mathrm{d}x, \tag{8.1}$$

for controls $u \in U_{\mathrm{ad}}^H$, and γ_1, $\gamma_2 > 0$.

Now, we show that our analysis for $n \leq 3$ applies to the corresponding FP ensemble optimal control problem subject to the following assumption on the PDF solution:

Assumption: We assume that there exists some small $\delta > 0$ such that the solution of our FP problem satisfies

$$\operatorname*{ess\,inf}_{(t,x)\in\Omega_T} p(t, x) > \delta \tag{8.2}$$

for all $u \in U_{\mathrm{ad}}^H$.

This assumption appears reasonable by considering an initial PDF $p_0 > \delta > 0$ and a bounded b, such that the drift acting on the stochastic process does not produce a region in Ω where the probability to find the state is below a certain threshold. In this case, (8.2) is preserved in the sense that $p(t) > \delta/\exp(t\|b\|_\infty)$ a.e. on Ω.

With this uniform positivity of p at hand, we can bound the minimizing sequence $(u^k) \subset U_{\mathrm{ad}}^H$ similarly to (5.6)

$$\Lambda\,\delta\|u^k\|^2_{H^1(\Omega_T)} \leqslant \int_{\Omega_T} \left(\frac{\gamma_1}{2}|u^k(t,x)|^2 + \frac{\gamma_2}{2}|\mathrm{D}u^k(t,x)|^2\right) p_k(t,x)\,\mathrm{d}t\,\mathrm{d}x \leqslant C + \hat{J}(u^k),$$

where $C = C(\Omega_T, \sigma, c, M, \alpha, \beta, \mathcal{T}, \|p_0\|_\infty)$. The weak lower semicontinuity follows with an application of Egorov's theorem, see Lemma 3.5 in [19]. Therefore, it is necessary to have $p_k \to p$ in $L^\infty(\Omega_T)$ and the non-negativity of the PDFs p_k, p. This proves the existence of an optimal control $u \in U_{\mathrm{ad}}^H$ for the objective functional J_1.

Next, we illustrate the first- and second-order analysis for the case J_1 with assumption (8.2). Clearly, all the properties of the control-to-state map G remain unchanged. A straightforward calculation with the product rule now shows that

$$\begin{aligned}
\hat{J}_1'(u)v &= \int_{\Omega_T} (\alpha\cdot v + \gamma\, u\cdot v + \gamma_2\,\mathrm{D}u\cdot\mathrm{D}v)\,p\,\mathrm{d}t\,\mathrm{d}x \\
&\quad + \int_{\Omega_T} \left(\mathcal{R}[u] + \frac{\gamma}{2}|u|^2 + \frac{\gamma_2}{2}|\mathrm{D}u|^2\right) z\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T} z(T)\,\mathrm{d}x \\
\hat{J}_1''(u)v^2 &= \int_{\Omega_T} 2\left(\alpha\cdot v + \gamma|v|^2 + \gamma_2|\mathrm{D}v|^2\right) p\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega_T} (\alpha\cdot v + \gamma\, u\cdot v + \gamma_2\,\mathrm{D}u\cdot\mathrm{D}v)\,z\,\mathrm{d}t\,\mathrm{d}x \\
&\quad + \int_{\Omega_T} \left(\mathcal{R}[u] + \frac{\gamma}{2}|u|^2 + \frac{\gamma_2}{2}|\mathrm{D}u|^2\right) w\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \mathcal{T} w(T)\,\mathrm{d}x.
\end{aligned}$$

where $p = G(u)$, $z = G'(u)v$ and $w = G''(u)v^2$ for $u \in U_{\mathrm{ad}}, v \in H^1(\Omega_T)^m$. Analogously to the previous section, we can show that $\hat{J}_1$ fulfils the regularity conditions (A1), (A2.1) and (A2.2).

Of course, the critical condition is (A2.3) since we replaced the H^1-norm of u with its expected-value. However, since $\gamma_1, \gamma_2 > 0$ and (8.2), for $v^k \rightharpoonup 0$ in $H^1(\Omega_T)^m$, we obtain for the crucial term of $\hat{J}''(u^k)(v^k)^2$ that we have for a subsequence

$$\Lambda\,\delta\|v^k\|^2_{H^1} \leqslant \int_{\Omega_T} \left(\frac{\gamma_1}{2}|v^k(t,x)|^2 + \frac{\gamma_2}{2}|\mathrm{D}v^k(t,x)|^2\right) p_k(t,x)\,\mathrm{d}t\,\mathrm{d}x.$$

This proves the sufficient second-order condition (A2) and completes the first- and second-order analysis for the expected-type cost functional J_1 with H^1-controls in U_{ad}^H.

In [15, 16], it is proposed to construct an approximate feedback control mechanism that requires to determine only time-dependent controls, whereas the dependence on x is chosen *a priori*. In particular, since many control mechanisms for SDE models have a linear or bilinear structure, it is reasonable to focus on the following controlled drift

$$B[u](t,x) := c(t,x) + u^0(t) + x \otimes u^1(t), \tag{8.3}$$

where c represents a given smooth vector field and $\otimes$ denotes the Hadamard product. Notice that this structure can also be motivated in the framework of first- and second-moment equations where it appears that u^1 enters in the control of the variance of the process, whereas its mean results controlled by both u^0 and u^1.

We remark that the drift (8.3) has the form of (2.2) with $n \times 2n$-matrix

$$M(t,x) = \begin{pmatrix} 1 & 0 & \dots & 0 & x_1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & x_2 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & x_n \end{pmatrix}.$$

In this case, we aim to minimize the following functional

$$J_2(p,u) = \int_{\Omega_T} \left(\mathcal{R}[u](t,x) + \frac{\gamma_1}{2}\,|u(t)|^2 + \frac{\gamma_2}{2}\,|\frac{\mathrm{d}}{\mathrm{d}t} u(t)|^2 \right) p(t,x)\,\mathrm{d}t\,\mathrm{d}x + \int_{\Omega} \mathcal{T}(x)\,p(T,x)\,\mathrm{d}x, \tag{8.4}$$

for $u(t) := (u^0(t), u^1(t)) \in \mathbb{R}^m$, $m = 2n$, from the admissible set

$$U_{\mathrm{ad}}^T := \left\{ u \in H^1(0,T)^m \,:\, u_i^{\min} \leqslant u_i \leqslant u_i^{\max}, \text{ a.e. on } \;]0,T[\,, i = 1,\dots,m \right\},$$

with given $u^{\min}, u^{\max} : [0,T] \to \mathbb{R}^m$ measurable and bounded such that $U_{\mathrm{ad}}^T \neq \emptyset$. We remark that in (8.4), the time derivative of the control is due to Brockett's concept of minimum attention control aim at penalizing large variation of the control, *cf.* [14].

For this setting, all the important properties of the control-to-state map corresponding to the drift (8.3) can be shown analogously to the previous sections. In particular, existence of optimal controls follows immediately: We pick a minimizing sequence $(u^k) \subset U_{\mathrm{ad}}^T$, show $u^k \rightharpoonup u$ in $H^1(0,T)$ and obtain $J(u^k, G(u^k)) > -\infty$ and $G(u^k) \to G(u)$ in $W(0,T)$ for a subsequence. Proving weak lower semicontinuity is easier, since all p_k, p are PDFs, and we have

$$\int_{\Omega_T} \left(\frac{\gamma_1}{2}|u(t)|^2 + \frac{\gamma_2}{2}|\mathrm{D}u(t)|^2 \right) p(t,x)\,\mathrm{d}t\,\mathrm{d}x = \int_0^T \left(\frac{\gamma_1}{2}|u(t)|^2 + \frac{\gamma_2}{2}|\mathrm{D}u(t)|^2 \right) \mathrm{d}t. \tag{8.5}$$

This proves the existence of an optimal control in U_{ad}^T.

Our second-order analysis can be applied analogously with the L^∞-convergences of the Fréchet-derivatives of $G(u^k)$ and using the trick of (8.5) to verify assumptions $(A1)$ and $(A2)$ from Theorem 7.2. This implies H^1-local uniqueness; see Corollary 7.6.

9. Conclusion

This work was devoted to the investigation of ensemble optimal control problems governed by a Fokker–Planck equation. These problems require the minimisation of objective functionals of probability type and aim at determining robust space-time control mechanisms for the ensemble of trajectories of the stochastic system defining the Fokker–Planck model. In this work, existence of optimal controls was proved and detailed analysis of first- and second-order optimality conditions characterizing these controls was presented.

Acknowledgements. The authors wish to express their gratitude to Maria Soledad Aronna, Fredi Tröltzsch and Daniel Wachsmuth for very helpful discussions. The authors also thank the two anonymous referee for their careful review of this paper and suggestions for improvement.

References

[1] R.A. Adams and J.J. Fournier, Sobolev Spaces. Elsevier (2003).

[2] M. Annunziato and A. Borzì, A Fokker–Planck control framework for multidimensional stochastic processes. *J. Comput. Appl. Math.* **237** (2013) 487–507.

[3] M. Annunziato and A. Borzì, A Fokker–Planck control framework for stochastic systems. *EMS Surv. Math. Sci.* **5** (2018) 65–98.
[4] M.S. Aronna and F. Tröltzsch, First and second order optimality conditions for the control of Fokker–Planck equations. *ESAIM: COCV* **27** (2021) 15.
[5] J. Bartsch, A. Borzì, F. Fanelli and S. Roy, A theoretical investigation of Brockett's ensemble optimal control problems. *Calc. Variat. Partial Differ. Equ.* **58** (2019) 1–34.
[6] J. Bartsch, A. Borzì, F. Fanelli and S. Roy, A numerical investigation of Brockett's ensemble optimal control problems. *Numer. Math.* **149** (2021) 1–42.
[7] J. Bartsch, G. Nastasi and A. Borzì, Optimal control of the Keilson-Storer caster equation in a Monte Carlo framework. *J. Comput. Theor. Transport* **50** (2021) 454–482.
[8] M. Bengfort, H. Malchow and F.M. Hilker, The Fokker–Planck law of diffusion and pattern formation in heterogeneous environments. *J. Math. Biol.* **73** (2016) 683–704.
[9] A. Bensoussan, Stochastic Control by Functional Analysis Methods. Elsevier (2011).
[10] A. Borzì, The Fokker–Planck framework in the modeling of pedestrians' motion, Chapter 6 in *Crowd Dynamics, Volume 2, Modeling and Simulation in Science, Engineering and Technology*, edited by L. Gibelli. Birkhäuser, Cham, Switzerland (2020) 111–131.
[11] T. Breiten, K. Kunisch and L. Pfeiffer, Control strategies for the Fokker–Planck equation. *ESAIM: COCV* **24** (2018) 741-763.
[12] T. Breitenbach and A. Borzì, The Pontryagin maximum principle for solving Fokker–Planck optimal control problems. *Comput. Optim. Appl.* **76** (2020) 499–533.
[13] T. Breitenbach and A. Borzì, A sequential quadratic Hamiltonian method for solving parabolic optimal control problems with discontinuous cost functionals. *J. Dyn. Control Syst.* **25** (2019) 403–435.
[14] R.W. Brockett, Minimum attention control, in *Proceedings of the 36th IEEE Conference on Decision and Control* 3 (1997) 2628–2632.
[15] R.W. Brockett, Optimal control of the Liouville equation. Proceedings of the International Conference on Complex Geometry and Related Fields. *AMS/IP Stud. Adv. Math.* **39** (2007) 23–35.
[16] R.W. Brockett, Notes on the control of the Liouville equation. Control of partial differential equations, Lecture Notes in Math. 2048. Springer, Heidelberg (2012) 101–129.
[17] B. Buonomo, D. Lacitignola and C. Vargas-De-León, Qualitative analysis and optimal control of an epidemic model with vaccination and treatment. *Math. Comput. Simul.* **100** (2014) 88–102.
[18] B. Buonomo, On the optimal vaccination strategies for horizontally and vertically transmitted infectious diseases. *J. Biolog. Syst.* **19** (2011) 263–279.
[19] E. Casas and F. Tröltzsch, Second order analysis for optimal control problems: improving results expected from abstract theory. *SIAM J. Optim.* **22** (2012) 261–279.
[20] E. Casas and F. Tröltzsch, Second order optimality conditions and their role in PDE control. *Jahresber. Dtsch. Math. Ver.* **117** (2015) 3–44.
[21] M. Chipot, Elements of Nonlinear Analysis. Springer Science & Business Media (2000).
[22] L.C. Evans,. Partial differential equations. *Graduate Stud. Math.* **19** (1998) 7.
[23] A. Fleig and R. Guglielmi, Optimal control of the Fokker–Planck equation with space-dependent controls. *J. Optim. Theory Appl.* **174** (2017) 408–427.
[24] D.A. Gomes and J. Saúde, Mean field games models – a brief survey. *Dyn. Games Appl.* **4** (2014) 110–154.
[25] P. Grandits, R.M. Kovacevic and V.M. Veliov, Optimal control and the value of information for a stochastic epidemiological SIS-model. *J. Math. Anal. Appl.* **476** (2019) 665–695.
[26] P. Han, Z. Chang and X. Meng, Asymptotic dynamics of a stochastic SIR epidemic system affected by mixed nonlinear incidence rates. *Complexity* (2020) 2020.
[27] J.M. Lasry and P.L. Lions, Mean field games. *Jpn. J. Math.* **2** (2007) 229–260.
[28] J.L. Lions and E. Magenes, Non-Homogeneous Boundary Value Problems and Applications (Vol. 1). Springer Science & Business Media, Berlin Heidelberg, (1972).
[29] H. Lou and J. Yong, Optimality conditions for semilinear elliptic equations with leading term containing controls. *SIAM J. Control Optim.* **48** (2009) 2366–2387.
[30] H. Lou and J. Yong, Second-order necessary conditions for optimal control of semilinear elliptic equations with leading term containing controls. *Math. Control Related Fields* **8** (2018) 57–88.
[31] X. Mao, Stochastic Differential Equations and Applications. Elsevier (2007).
[32] A. Rösch and D. Wachsmuth, Numerical verification of optimality conditions. *SIAM J. Control Optim.* **47** (2008) 2557–2581.
[33] H. Risken, Fokker–Planck Equation. Springer, Berlin, Heidelberg (1996).
[34] S. Roy, M. Annunziato, A. Borzì and C. Klingenberg, A Fokker–Planck approach to control collective motion. *Comput. Optim. Appl.* **69** (2018) 423–459.
[35] S. Roy, M. Annunziato and A. Borzì, A Fokker–Planck feedback control-constrained approach for modelling crowd motion. *J. Comput. Theor. Transport* **45** (2016) 442–458.
[36] S. Roy, A. Borzì and A. Habbal, Pedestrian motion modelled by Fokker–Planck Nash games. *Royal Soc. Open Sci.* **4** (2017) 170648.
[37] M. Schienbein and H. Gruler, Langevin equation, Fokker–Planck equation and cell migration. *Bull. Math. Biol.* **55** (1993) 585–608.

[38] R.K. Tagiyev, Optimal control by the coefficients of a parabolic equation. *Trans. Acad. Sci. Azerb. Ser. Phys.-Tech. Math. Sci.* **24** (2004) 247–256.
[39] R.K. Tagiyev, On optimal control of the hyperbolic equation coefficients. *Autom. Remote Control* **73** (2012) 1145–1155.
[40] D. Wachsmuth, The regularity of the positive part of functions in $L^2(I; H^1(\Omega)) \cap H^1(I; H^1(\Omega)^*)$ with applications to parabolic equations. *Comment. Math. Univers. Carolinae* **57** (2016) 327–332.
[41] Z. Wu, J. Yin and C. Wang, Elliptic & Parabolic Equations. World Scientific (2006).
[42] Y. Zhang, Y. Li, Q. Zhang and A. Li, Behavior of a stochastic SIR epidemic model with saturated incidence and vaccination rules. *Physica A* **501** (2018) 178–187.